Investigations into the regulation of sister chromatid cohesion

Katarzyna Kozyrska

Trinity College

Department of Biochemistry

University of Oxford

Thesis submitted for the degree of Master of Science (by Research)

Trinity Term 2014

Acknowledgements

I must express my gratitude to the Wellcome Trust for providing me with such a generous stipend. It was thanks to this that I could act on my hatred of carpets and non-mixer taps and live like a true continental European.

I would like to thank the Nasmyth lab for hosting me during my journey into the murky waters of biochemistry. I would especially like to extend my gratitude to Frédéric Beckouët for his critical reading of all my last-minute scientific concoctions and for showing me how bad I am at wrestling. I would also like to thank Lana Strmecki for her miracle work in extracting signatures from elusive characters and for letting me spoil Spike over the summer.

I am infinitely grateful to my twin, Naomi 'Norman' Petela, whom I could always rely on to indulge in our shared love of good food, bad TV, and C_2H_5OH (of any quality). I am also thankful to both Naomi 'Norman' Petela and Peter West for pushing me to explore foreign territories such as enjoying chilli peppers and not being angry at board games. I am yet to master either of these.

I would like to give my thanks to Jean Metson and Martin Houlard for persuading/bullying/blackmailing me into doing my first half marathon. I now appreciate that dignity is of little importance when fresh underwear is concerned after a race. I would also like to thank Jean Metson for introducing me to the masochistic appeal of camping.

Finally, I would like to thank all the non-Nasmyth friends I have been fortunate enough to make during my stay in Oxford: Nadine Muschalik, Violet (Zhe) Feng, Metta Pratt, Jenny Richens, Emilie Bauwens, Cat 'Vincent' Vicente, and many others. It truly would have been a bleak time without the help and support of each and every one of you. Also, I would probably be in prison by now, had you not been there to recognise my non-PC humour for what it was/is.

Abstract

Cohesin, a ring-shaped complex composed of four subunits, is an essential player involved in timely and accurate segregation of genetic material at mitosis and meiosis. Cohesin performs a highly conserved role by topologically embracing sister chromatids until their concerted disjunction at anaphase, when the α-kleisin subunit of the ring is irreversibly cleaved by separase. A central part of the cohesin cycle is its cleavage-independent removal from chromatin through the action of the releasing complex.

In more complex eukaryotes, stabilisation of cohesin on DNA depends on the presence of sororin, a small functionally conserved protein whose association with cohesin is thought to counteract the releasing reaction. The first part of this thesis aimed to address the discrepancy of a sororin orthologue never having been identified in *S. cerevisiae*. This was undertaken using an imaging-based approach to screen a subset of the yeast genome for proteins with cohesin-like localisation. The screen was streamlined such that candidate genes had periodically cycling mRNA transcripts, similarly to cohesin subunits. No novel cohesin-associated proteins could be identified, although this could be in part due to technical limitations of the screen.

The precise molecular details of cohesin's function and regulation remain poorly understood. A conclusive way to address many questions about cohesion as a whole would be to set up an *in vitro* assay capable of reconstructing cohesin loading, cohesion establishment, and the releasing reaction. The second part of this thesis aimed to establish a reliable purification protocol for the *S. cerevisiae* cohesin complex using a bacterial expression system for use in such biochemical assays. While purification of all components of the cohesin trimer (Smc1, Smc3, Scc1) was achieved successfully, the complex could not be confirmed to assemble functionally, as determined by chemical crosslinking of the Smc3/Scc1 interface.

Contributions

The work presented in this thesis was funded by the Wellcome Trust.

Parts of this thesis were completed thanks to data obtained through collaboration with other researchers, as follows:

The *S. cerevisiae* dynamic transcriptome analysis (DTA), which generated the list of genes with periodically cycling mRNAs used in the screen for 'yeast sororin', was performed by Prof Patrick Cramer (Ludwig Maximilian University of Munich, Germany).

The mass spectrometry analysis used to identify protein species following affinity purification was performed by Dr Benjamin Thomas (Central Proteomics Facility, Sir William Dunn School of Pathology, University of Oxford).

List of abbreviations

°C	Degrees Celsius
2×TY	2 times tryptone and yeast extract
6×His (tag)	6-histidine affinity tag
A	Amperes
Å	Ångströms
APC/C	Anaphase-Promoting Complex/Cyclosome
ATP	Adenosine 5'-triphosphate
BMOE	Bismaleimidoethane
bp	Base pair(s)
BSA	Bovine Serum Albumin
C- (terminal)	Carboxy-
Cam	Chloramphenicol
CDC (Cdc)	Cell Division Cycle
CEN	Core centromeric sequence
ChIP	Chromatin immunoprecipitation
co-IP	Co-immunoprecipitation
Cohesin tetramer	Smc1, Smc3, Scc1, Scc3
Cohesin trimer	Smc1, Smc3, Scc1
CV	Column Volume
D. melanogaster	*Drosophila melanogaster*
dBBr	Dibromobimane

ddH_20	Double-distilled water
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic Acid
dNTPs	Deoxyribonucleotide Triphosphates
DTA	Dynamic Transcriptome Analysis
DTT	Dithiothreitol
E. coli	*Escherichia coli*
ECL	Enhanced Chemiluminescence
ECO (Eco)	Establishment of Cohesion
EDTA	Ethylenediaminetetraacetic Acid
EM	Electron Microscopy
Fig.	Figure
FKBP12	FK506 binding protein 1A, 12 kDa
Frb	FKBP12-rapamycin binding
FRET	Förster Resonance Energy Transfer
FT	Flow-through
g	Gravitational force
G1 (phase)	First gap phase
G2 (phase)	Second gap phase
GFP	Green Fluorescent Protein
HCl	Hydrochloric acid
HF (restriction enzyme)	High Fidelity
His	Histidine
hr	Hour(s)
HRP	Horseradish Peroxidase

IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
kb	Thousand base pairs
kDa	Kilo Daltons
kpsi	Thousand Pounds per Square Inch
LB	Lysogeny broth
LDS	Lithium Dodecyl Sulphate
M	Molar
M (phase)	Mitotic phase/mitosis
mA	Milliamperes
mg	Milligrams
min	Minute(s)
ml	Millilitres
mM	Millimolar
mRNA	Messenger Ribonucleic Acid
MT	Microtubule
MTW (Mtw)	Mis Twelve-like
MW	Molecular Weight/mass
N- (terminal)	Amino-
NaCl	Sodium chloride
NBD	Nucleotide Binding Domain
ng	Nanograms
nM	Nanomolar
OD_{600}	Optical Density at 600 nm

oligo	DNA oligonucleotide
ORF	Open Reading Frame
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline + 0.05% Tween 20
PCR	Polymerase Chain Reaction
pI	Isoelectric point
PMSF	Phenylmethylsulphonylfluoride
PVDF	Polyvinylidene Difluoride
RBS	Ribosome Binding Sequence
RFP	Red Fluorescent Protein
RNA	Ribonucleic Acid
rpm	Revolutions Per Minute
RT	Room Temperature
S (phase)	Synthesis phase
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
SAC	Spindle Assembly Checkpoint
SCC (Scc)	Sister Chromatid Cohesion
SDM	Site-Directed Mutagenesis
SDS	Sodium Dodecyl Sulphate
sec	Second(s)
SMC (SMC, Smc)	Structural Maintenance of Chromosomes
SPB	Spindle Pole Body
T_m	Melting temperature

Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer Ribonucleic Acid
UV	Ultraviolet
V	Volts
v/v	Volume/volume
w/v	Mass/volume
WAPL (Wapl)	Wings-Apart Like
WCE	Whole Cell Extract
WT	Wild-Type
X. laevis	*Xenopus laevis*
YEP	Yeast Extract Peptone
YPD	Yeast Extract Peptone + 2% Dextrose
µg	Micrograms
µl	Microlitres
µM	Micromolar

Table of contents

CHAPTER I

Introduction

CHAPTER I: Introduction

Cell division is arguably the single most important function a living organism must be able to perform. Most unicellular species rely on mitosis for reproduction, while more complex multicellular organisms require it for growth and tissue damage repair. Being so crucial to survival, cell division has evolved to be a robust, highly conserved process ensuring that genomic DNA is replicated accurately and that the two copies are separated into two genetically identical daughter cells. That said, when problems in cell cycle regulation do occur, they can have severe consequences, an example being cancer in multicellular organisms.

The cell cycle can be divided into four discrete stages: first gap (G1), synthesis (S), second gap (G2), and mitotic (M) phases (Figure 1). Progression from one stage to the next is regulated by a series of molecular switches, which ensure timely and unidirectional progression through the four phases. In G1, cycling cells grow prior to making the commitment to replicate their genetic material, which occurs in S phase. In G2, completion of replication is verified before cells can commit to mitosis, which itself is comprised of four stages: prophase, metaphase, anaphase, and telophase (Morgan, 2007). In prophase, the replicated genome is condensed to reveal pairs of sister chromatids microscopically visible as distinct 'rods' (Sumner, 1991). To ensure timely and accurate segregation of genetic material, the duplicated chromosomes must then be bi-oriented at the

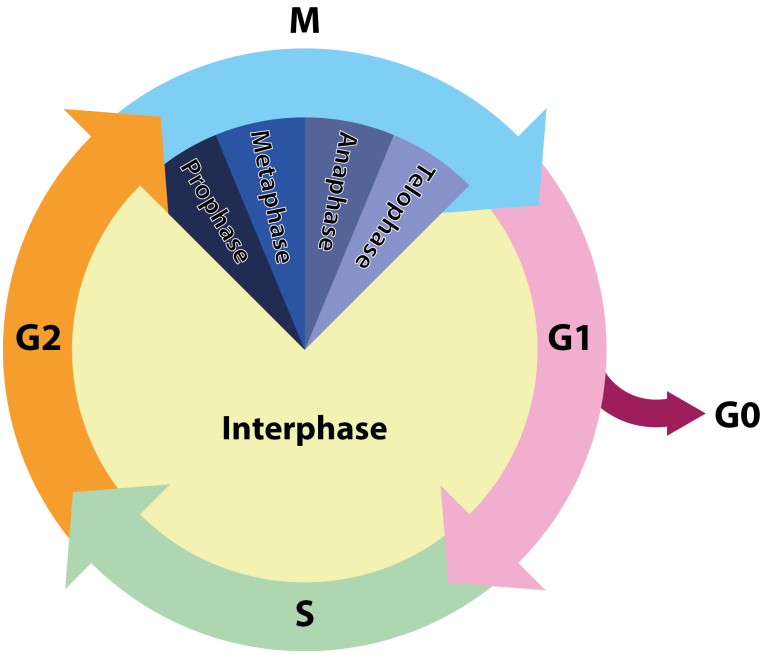

Figure 1. The cell cycle.

The cell cycle is a highly regulated, unidirectional progression with the aim of producing two genetically identical daughter cells from a single mother cell. It can be divided into four stages: first gap (G1) phase where cells prepare for division; synthesis (S) phase where cells replicate their entire genome; second gap (G2) phase where DNA replication is verified and any damage repaired; and mitotic (M) phase where chromosomes are condensed and segregated. M phase can be further subdivided into prophase, metaphase, anaphase, and telophase. If a cell does not commit to DNA replication, it may exit the cell cycle and become quiescent (G0).

[Adapted from S. Dixon (2013). DPhil thesis. University of Oxford, UK.]

metaphase plate in a process that relies on tension generated when the pulling force of microtubules (MTs) on each sister chromatid opposes the cohesion between them. Inter-chromatid cohesion is generated through two mechanisms: DNA catenation created as a consequence of replication, and the cohesin complex that physically tethers sister chromatids together.

The spindle assembly checkpoint (SAC) inhibits anaphase progression until each kinetochore pair has been captured by MTs emanating from opposite poles of the cell, generating tension (Foley and Kapoor, 2013). Tension across the

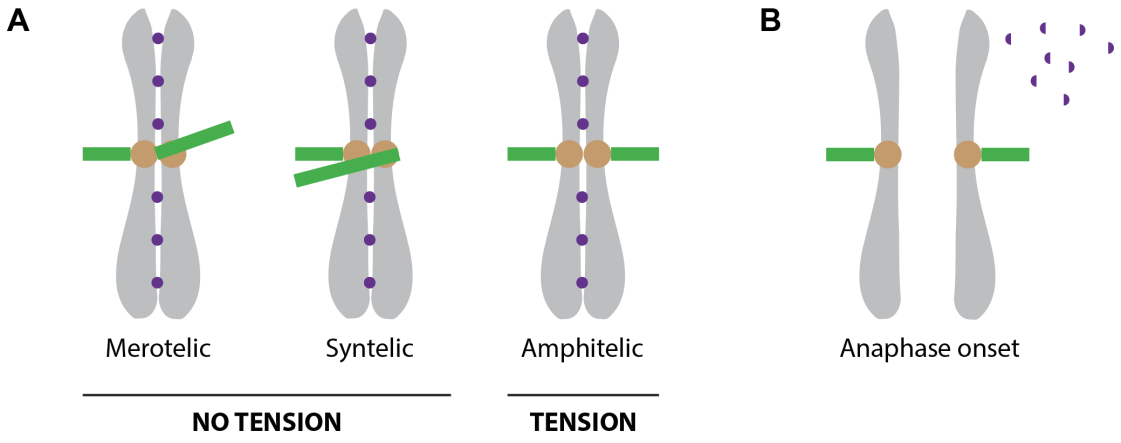

Figure 2. Chromosome bi-orientation is key to mitotic progression.

(A) Incorrect chromosome/microtubule (MT) interactions such as merotelic or syntelic attachments do not generate tension across the kinetochore. Tension can only be generated as a result of amphitelic MT attachments. [DNA has been represented in grey, kinetochores in brown, MTs in green, and cohesin as purple dots.]
(B) Proteolytic cleavage of cohesin at anaphase releases the sister chromatids such that they can be pulled apart by the mitotic spindle into the two daughter cells.

kinetochore is only generated in the case of amphitelic MT attachments (Figure 2A); erroneous MT/kinetochore interactions are destabilised by the Aurora B kinase in an error-correction mechanism (Liu and Lampson, 2009). Once all chromosomes have been bi-oriented at the metaphase plate, the SAC is silenced to activate the anaphase-promoting complex/cyclosome and its regulator Cdc20 (APC/C^{Cdc20}). APC/C^{Cdc20} prevents inhibitory phosphorylation of the protease separase and targets its negative regulator, securin, for degradation (Musacchio and Salmon, 2007). Cohesion is then dissolved at anaphase by proteolytic cleavage of cohesin by the now-active separase, allowing the two sets of chromosomes to be pulled apart into the two daughter cells during telophase (Figure 2B; Ciosk et al. 1998). Mitotic exit is promoted by the APC/C^{Cdh1} complex, which targets mitotic determinants for degradation (Sullivan and Morgan, 2007).

Cohesin is a complex composed of four core protein subunits: a heterodimer of structural maintenance of chromosomes (SMC) proteins (Smc1 and Smc3), an α-kleisin subunit (Scc1 in *Saccharomyces cerevisiae; S. cerevisiae*), and Scc3. SMC proteins adopt a highly conserved tertiary rod-like structure: a central kink in the primary sequence induces the formation of a 'hinge' dimerisation domain at one end of the SMC 'rod' connected by a coiled-coil to a nuclear binding domain (NBD ATPase 'head') formed as a consequence of the two termini coming together (Figures 3A-C; Haering et al. 2002; Hirano & Hirano 2002). Smc1 and Smc3 form a very stable heterodimer via interactions between their respective 'hinge' domains, enabling their co-purification and visualisation by electron microscopy (EM) as V-shaped molecules (Anderson et al., 2002; Haering et al., 2002). Association of the C and N termini of Scc1 with Smc1 and Smc3 NBDs, respectively, forms a giant proteinaceous ring capable of embracing sister chromatids (Gruber et al., 2003; Haering et al., 2008). Although all three interfaces of cohesin have been crystallised, the structure of the core subunits and, consequently, that of the tetramer remains unsolved (Gligoris et al., 2014; Haering et al., 2004; Kurze et al., 2011). The final member of the cohesin tetramer, Scc3, interacts with the ring in a stoichiometrical fashion via Scc1 (Figure 3D; Haering et al. 2002). Scc3 has been shown to be essential for both loading and maintenance of cohesin on chromatin, which makes it difficult to separate its role in maintaining complex integrity and any cohesin cycle-specific function(s) (Toth et al., 1999).

The ring-like architecture of cohesin, combined with the fact that its cleavage induces sister chromatid separation, led to the idea that cohesin topologically

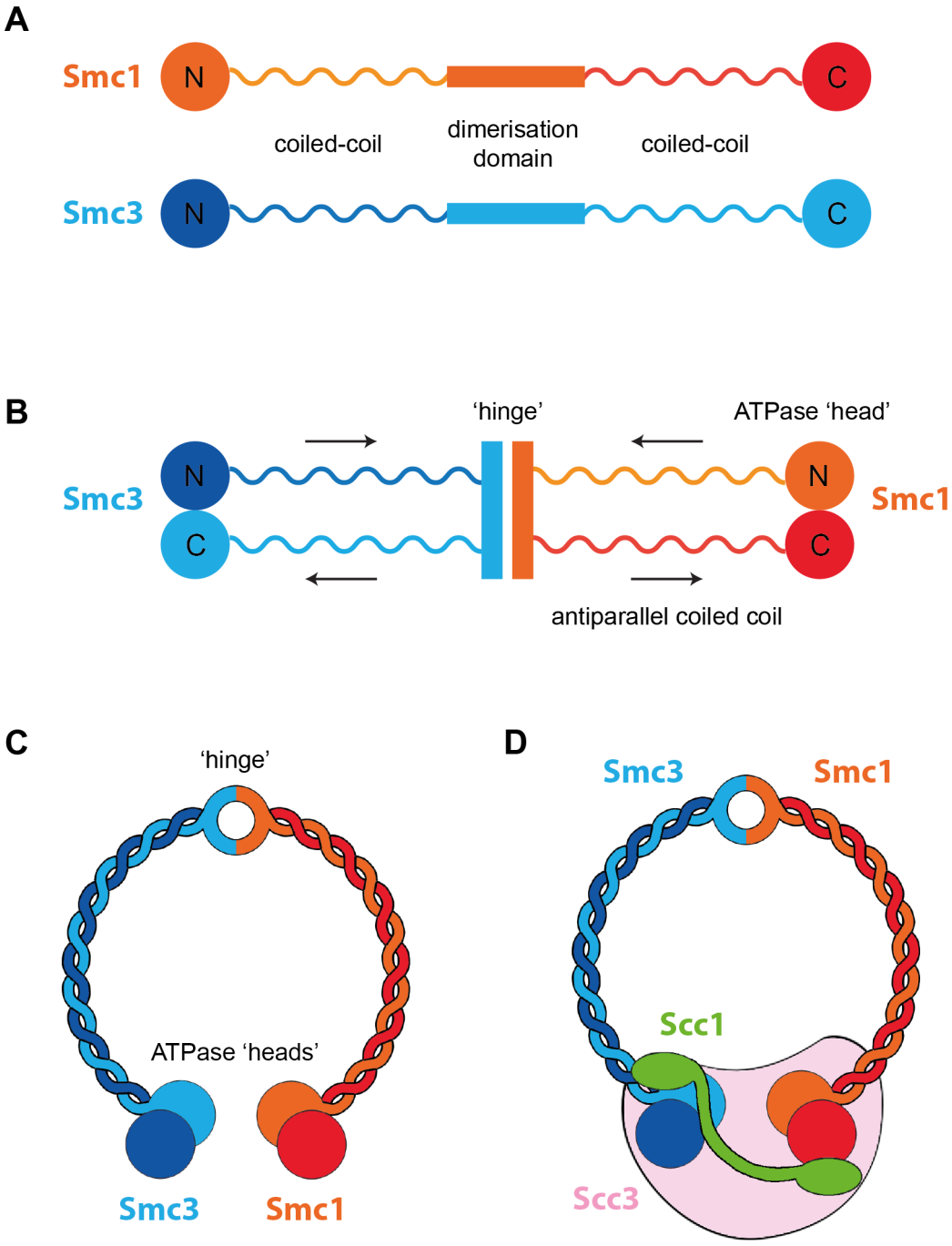

Figure 3. The architecture of the Smc1/Smc3 heterodimer and the cohesin tetramer.

(A) Schematic representations of functional domains within structural maintenance of chromosomes (SMC) proteins Smc1 and Smc3.
(B-C) The central dimerisation interface in each SMC protein induces the formation of a fold in the middle of the peptide and brings the two termini together, resulting in a functional ATPase 'head' domain. Dimerisation domains of Smc1 and Smc3 then come together to form an interface termed the 'hinge.' Smc1/Smc3 heterodimers adopt a characteristic V-shaped conformation, as observed by electron microscopy.
(D) The NBDs of the Smc1/Smc3 heterodimer are bridged by Scc1 to complete the cohesin ring. Scc3 associates via Scc1 and is considered a part of the core cohesin complex.

[Adapted from W. Upcher (2012). DPhil thesis. University of Oxford, UK.]

embraces sister DNAs. Introduction of chemically cross-linkable cysteine pairs at Smc1/Smc3 and Smc1/Scc1 interfaces with simultaneous expression of a fusion Smc3/Scc1 protein creates a version of cohesin that can be induced to form one continuous, covalently-linked proteinaceous circle. Upon replication of yeast minichromosomes, the resulting cohesed dimers have been shown to remain dimeric even after denaturation of associated proteins (Haering et al., 2008). These data are consistent with the model whereby the cohesin ring topologically entraps replicated sister DNAs until anaphase, resulting in sister chromatid cohesion (Farcas et al., 2011; Gligoris et al., 2014; Haering et al., 2008; Nasmyth and Haering, 2005; Oliveira et al., 2010; Uhlmann et al., 1999, 2000).

A corollary of cohesin's topological entrapment of DNA is that one of the ring's interfaces must open to allow entry of DNA into the ring lumen. In *S. cerevisiae*, cohesin is loaded onto chromatin during late G1 by the kollerin complex (Scc2, Scc4), and this depends on ATP hydrolysis by the SMC NBDs (Arumugam et al., 2003; Ciosk et al., 2000; Hu et al., 2011). Insertion of human FKBP12 and Frb domains into the yeast Smc1 and Smc3 hinges creates a situation whereby the cohesin Smc1/Smc3 interface can be conditionally 'locked,' which abolishes the ability of cohesin to load onto DNA but has little effect on cohesion maintenance (Gruber et al., 2006). This evidence strongly suggests that cohesin loading occurs via transient opening of the Smc1/Smc3 hinge interface to enable DNA entry into the ring lumen (Figure 4A; Gruber et al. 2006). That said, the hinge modifications used in these experiments are rather bulky and may have had an indirect effect on cohesin's architecture, thereby not unequivocally proving the identity of its DNA 'entry gate.'

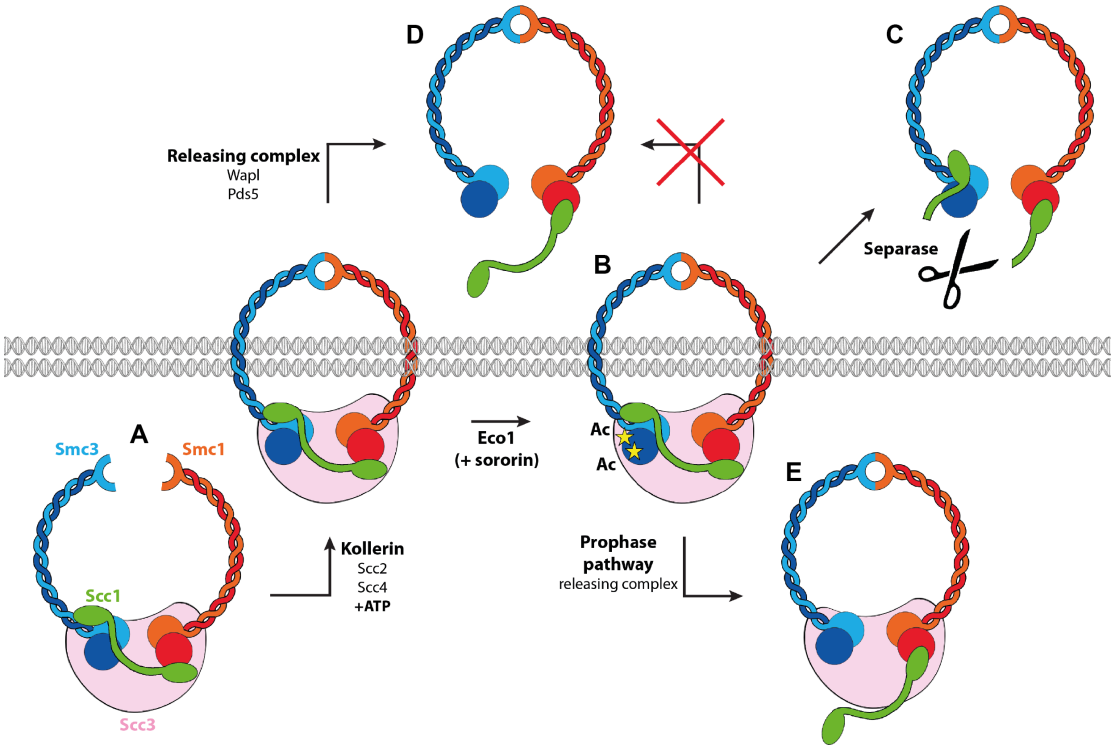

Figure 4. Schematic representation of the cohesin cycle.

(A) Cohesin is loaded onto DNA in late G1 via transient opening of the Smc1/Smc3 'hinge,' mediated by the kollerin complex, in the presence of ATP.
(B) Acetylation of the Smc3 head domain by Eco1 during DNA replication, and subsequent recruitment of sororin in complex eukaryotes, 'locks' the cohesin ring against releasing activity.
(C) At anaphase, Scc1 is cleaved by separase, allowing for sister chromatid segregation.
(D) Prior to cohesion establishment, the cohesin ring is removed in a cleavage-independent manner from chromatin by the releasing complex, which promotes opening of the Smc3/Scc1 interface.
(E) In complex eukaryotes, a large proportion of cohesin bound on chromosome arms is removed in a process dependent on the releasing complex after cohesion establishment (prophase pathway). This reaction is counteracted by sororin, which competes with Wapl to stabilise the cohesin ring on DNA.

Cohesin becomes stably bound to DNA (i.e. cohesion is established) during DNA replication, when two conserved lysine residues in the Smc3 NBD are acetylated by the acetyltransferase Eco1 (Figure 4B; Ivanov et al. 2002; Heidinger-Pauli et al. 2010; Ben-Shahar et al. 2008; Ünal et al. 2008). Cohesion is maintained until anaphase, when the protease separase cleaves Scc1 to open the cohesin ring and thus removes it from DNA. This irreversibly dissolves sister chromatid cohesion

and allows the two sister chromatids to be pulled apart by the mitotic spindle into separate daughter cells (Figure 4C; Uhlmann et al. 1999).

There is a second, separase-independent process that removes cohesin rings from DNA (Figure 4D). Cohesin dissociation from chromatin can be detected throughout the mammalian cell cycle but is noticeably increased in prophase (Gerlich et al., 2006). This is when a large proportion of cohesin is removed from chromosome arms (a process termed the 'prophase pathway') and remains in the nucleoplasm as soluble, intact rings until reloading in telophase (Figure 4E; Sumara et al., 2000; Tedeschi et al., 2013; Waizenegger et al., 2000). Disruption of the prophase pathway can be induced by depletion of Wapl or Pds5 from *Xenopus laevis* (*X. laevis*) egg extracts and results in severe defects in sister chromatid resolution, demonstrating that cohesin's release from DNA is dependent on these two proteins (Gandhi et al., 2006; Kueng et al., 2006; Shintomi and Hirano, 2009). Wapl and Pds5 are thus referred to as the 'releasing complex.' Despite the lack of a defined prophase pathway in *S. cerevisiae*, the releasing activity of Wapl and Pds5 is conserved: photobleaching data indicate that cohesin associates dynamically with yeast chromatin prior to DNA replication in a process depending on Wapl (Chan et al. 2012).

As before, the ring's topological interaction with DNA dictates that one of its interfaces must open for cohesin's releasing complex-mediated dissociation from chromatin. While induction of Wapl expression in wild-type cells at G2/M removes DNA-bound cohesin in a cleavage-independent manner, cohesin

remains on chromatin in the presence of the Smc3/Scc1 fusion. This indicates that Wapl acts to displace cohesin by transiently opening the Smc3/Scc1 interface (Chan et al., 2012; Eichinger et al., 2013). That being said, expression of the Smc3/Scc1 fusion may have affected the complex indirectly (e.g. by altering the topology of another interface) and, hence, this approach does not provide definite proof of the identity of cohesin's DNA 'exit gate.'

In order to perform its 'cohesive' role, cohesin's association with chromatin switches from dynamic in late G1 to stable upon DNA replication, and remains stable up until cohesin cleavage in anaphase. This 'stabilisation' of cohesin's binding occurs in conjunction with Eco1-dependent acetylation of two adjacent lysine residues in the Smc3 head domain (Ben-Shahar et al., 2008; Rowland et al., 2009; Ünal et al., 2008). Induction of Wapl expression in *S. cerevisiae* at G2/M induces cohesion defects in *eco1Δ* cells, which indicates that Smc3 acetylation prevents the releasing reaction. Expression of the fusion Smc3/Scc1 protein bypasses the requirement for the essential Eco1 protein (Chan et al., 2012). This mimicking of Eco1 function by the Smc3/Scc1 fusion protein indicates that the primary role of Smc3 acetylation is to 'lock' the Smc3/Scc1 interface against the releasing activity.

Although the molecular details associated with Smc3 acetylation are yet to be determined, evidence from mammalian cells suggests that this modification recruits sororin, a small functionally conserved protein that is essential for the maintenance of cohesion (Rankin et al., 2005; Schmitz et al., 2007; Zhang and Pati, 2012). Upon acetylation of Smc3, sororin is recruited to chromatin-bound

cohesin complexes, where it is thought to displace Wapl from its binding sites on Pds5 in a competitive manner (Lafont et al., 2010; Nishiyama et al., 2010; Whelan et al., 2012). Although Wapl is not completely ejected from the complex, sororin's action is believed to counteract the releasing activity of the Wapl/Pds5 dimer, hence stabilising cohesin rings on the DNA. Despite being present and essential in all vertebrates and many invertebrates, a sororin orthologue are yet to be identified in budding yeast (Nishiyama et al., 2010; Zhang and Pati, 2012). A streamlined screen performed in this thesis attempted to address this discrepancy.

Wapl's association with cohesin depends on Pds5, although the former also interacts with Scc3 and possibly with other core cohesin subunits (Rowland et al., 2009; Shintomi and Hirano, 2009). Pds5, in turn, binds Scc1 and seems to play three rather distinct roles: it is necessary for promoting Smc3 acetylation during cohesion establishment, for cohesion maintenance, and for cleavage-independent cohesin release (Chan et al., 2013; Hartman et al., 2000; Panizza et al., 2000; Shintomi and Hirano, 2009).

Very little is known about the molecular details of the releasing mechanism, especially in complex eukaryotes where the releasing complex plays at least three essential roles. In addition to the discussed roles of the releasing complex in cohesion establishment and the prophase pathway (in vertebrates), budding yeast Wapl has been recently shown to prevent excessive condensation of chromosomes (Lopez-Serra et al., 2013). Given its involvement in several processes critical to the stability and organisation of the genome, the releasing

complex is a multifaceted target for investigation. By dissecting its mechanism in a simpler organism, such as *S. cerevisiae*, we can begin to understand the inner workings of the releasing reaction. What is known so far are the key players (Wapl, Pds5) and at least in some part how they interact. The next landmark would be to determine the fundamental requirements for releasing activity: Which interface(s) needs to open? Do all four core cohesin subunits need to be present? Is ATP hydrolysis required?

In order to address these questions, the latter half of this thesis describes a purification strategy to generate a recombinant version of the yeast cohesin complex for subsequent use in biochemical assays. The ultimate goal was to use the isolated cohesin complex in an *in vitro* DNA loading system to subsequently deduce what the minimum constituents for the releasing reaction are. To achieve this, the yeast cohesin tetramer (Smc1, Smc3, Scc1, Scc3) and its regulatory subunits (Scc2, Scc4, Pds5, Wapl) would need to be purified and the complex loaded onto plasmid DNA *in vitro*. Topological cohesin/DNA interaction would then be confirmed by determining whether cohesin immunoprecipitation can recover high salt wash-resistant DNA. Finally, the proportion of DNA released from cohesin rings after supplementing the mix with releasing complex components would be quantified. The logic behind this approach stemmed from a recently published *in vitro* assay for cohesin loading using *Schizosaccharomyces pombe* (*S. pombe*) proteins (Murayama and Uhlmann, 2014). While the assay itself was beyond the scope of this thesis, the recombinant complex (once confirmed that it assembles functionally) could nonetheless be used in multiple investigations that further our understanding of cohesion.

Thesis objectives

The primary objective of this thesis was to identify a sororin orthologue, or other novel cohesin regulatory proteins, in *S. cerevisiae* using an imaging-based screen of a subset of budding yeast genes. Dynamic transcriptome analysis data obtained by Prof Patrick Cramer's group (Ludwig Maximilian University of Munich, Germany) were used to narrow down the yeast genome to candidate genes whose mRNA transcription profiles behaved similarly to that of cohesin subunits. Namely, the mRNA of candidate genes was transcribed in a cell cycle-dependent manner and was of a similar periodicity to that of known cohesin genes.

The secondary objective of this thesis was to reconstitute the core cohesin tetramer (Smc1, Smc3, Scc1, Scc3) for use in various biochemical assays. An *in vitro* assay reproducing the conditions necessary for the loading reaction, cohesion establishment, and the releasing reaction would be the most indisputable and elegant way to address many questions about the molecular details of cohesin's function. Although the assay itself was beyond the scope of this project, the work presented in this thesis aimed to set up a reliable method for expression of a recombinant, functional cohesin complex.

CHAPTER II

A streamlined screen of the *S. cerevisiae* genome for ‘yeast sororin’ and other regulators of cohesin

CHAPTER II: A streamlined screen of the *S. cerevisiae* genome for 'yeast sororin' and other regulators of cohesin

Introduction

In budding yeast, the loading of cohesin onto chromatin occurs in late G1, following re-synthesis of Scc1, and this process is dependent on three factors: the Scc2/Scc4 complex ('kollerin'), functional kinetochores, and ATP binding and hydrolysis by the SMCs (Arumugam et al., 2003; Ciosk et al., 2000; Weitzer et al., 2003). Chromatin immunoprecipitation (ChIP) studies have shown that cohesin is enriched in a distinct pattern across chromosomes: in pericentromeric regions i.e. 20-50 kb stretches on either side of the centromere (*CEN*), and at sites spaced at ~11 kb intervals between convergent genes (Blat and Kleckner, 1999; Glynn et al., 2004; Laloraya et al., 2000; Megee et al., 1999; Tanaka et al., 1999). This distribution is believed to be a consequence of cohesin being loaded predominantly at centromeres and then translocated to its final residence sites. Cohesin's migration from its loading sites most likely occurs by sliding along chromatin fibres as a result of pushing forces generated by transcribing polymerases (Lengronne et al., 2004). That said, recent work from the lab has indicated that there may be two distinct pathways for cohesin loading. While loss of Scc4 function abolishes cohesin loading, a point mutation in Smc1 has the ability to restore loading at arm loci but, importantly, not at centromeres in a

scc4Δ scenario. This suggests loading at chromosome arms may occur via a different mechanism (Dr S. Dixon, personal communication).

Cohesin is associated with yeast chromosomes from late G1 until the metaphase/anaphase transition, when Scc1 is cleaved and degraded prior to its re-synthesis in the next G1. When visualised by fluorescence microscopy between G1 and anaphase, all core cohesin subunits display a characteristic distribution within the budding yeast nucleus. In small G1 phase cells, this appears as an enrichment of fluorescent cohesin signal that co-localises with a single centromere focus (as marked by the fluorescently labelled kinetochore protein Mis Twelve-like 1, Mtw1) at the edge of the nucleus, near the spindle pole body (SPB; Figure 5A; Yeh et al. 2008). In large G2/M phase cells, cohesin's fluorescent signal has a circular cross-section and co-localises with two centromere foci at either end of the cylinder (Figure 5B and Chan et al. 2012). This characteristic cylindrical pattern is termed the 'cohesin barrel' and can be observed for most cohesin-associated proteins, although it is less distinct for Wapl, for example.

The distinctive 'cohesin barrel' is suggested to be a consequence of pericentromeric chromatin organisation by the cohesin complex. While chromosome arms in budding yeast remain closely linked until anaphase, sister centromeres are flexible and fluctuate relative to one another most probably because they have been captured by kinetochore MTs. Cohesin rings have been postulated to arrange centromeric DNA in a cruciform manner, which would explain how centromeres remain linked even when they are a measurable

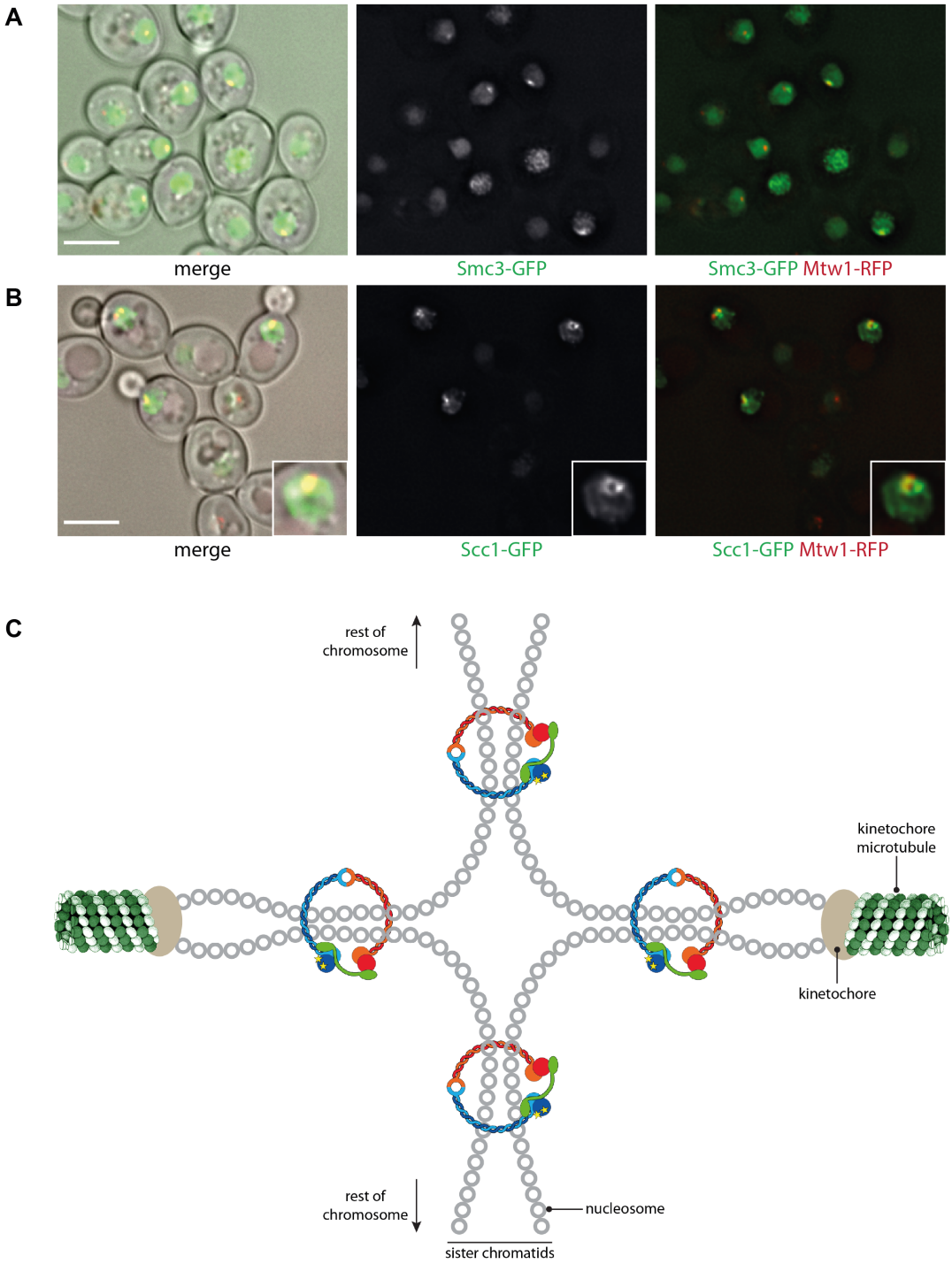

Figure 5. GFP fusions of cohesin subunits form the characteristic 'cohesin barrel' around the yeast centromere.

(A) In haploid G1 cells, Smc3-GFP showed an nuclear localisation characteristic of cohesin subunits: most of the fluorescent signal was localised at the nuclear edge and co-localised with a single Mtw1-RFP dot, which is a kinetochore marker. [All scale bars represent 5 µm.]
(B) In homozygous, diploid, G2/M cells, Scc1-GFP formed the characteristic 'cohesin barrel', a cylindrical structure that can be seen to co-localise with two Mtw1-RFP foci. The inset demonstrates a 2× magnification of the 'cohesin barrel' in transverse view.
(C) Schematic representation of the current model for cohesin's distribution on pericentric chromatin. [Adapted from Yeh et al. 2008.]

distance apart prior to anaphase (Figure 5C). Such arrangement of cohesin on all 32 replicated yeast centromeres around the mitotic spindle is proposed to generate a cylindrical structure, the 'cohesin barrel', on a macromolecular level (Yeh et al., 2008).

The first aim of this thesis was to address the discrepancy of a sororin orthologue not having been identified in budding yeast, despite the protein's highly conserved function in more complex eukaryotes. Since sororin's role in cohesion establishment is essential, determining whether *S. cerevisiae* possesses an orthologue would open up new important avenues of research into the regulation of cohesion in different organisms. If sororin is present in yeast, does it play the same role during the cohesin cycle? If it is absent, how does budding yeast achieve robust cohesion? How are the equivalent cohesin 'locking' changes induced? To clarify this inconsistency, a streamlined screen of the yeast genome was performed for 'yeast sororin' and any other novel cohesin-associated proteins.

A streamlined screen of the yeast GFP-fusion library to identify 'yeast sororin' and other novel cohesin regulators

Despite its conserved function, sororin's protein sequence has no recognisable structural domains or motifs, so traditional bioinformatics approached have yet to identify a sororin orthologue in less complex eukaryotes, including budding yeast (Wu et al., 2011). Although sororin's function in promoting the stable

binding of cohesin to chromatin has been recognised in many organisms, little is known about the molecular details of this process. It is therefore difficult to design a 'functional' screen for a sororin orthologue without a defined idea of what precise phenotype is being sought. One could envisage that a traditional forward genetic screen for strains defective in cohesion establishment, for example, would primarily yield hits in *ECO1* and *SMC3* rather than any novel regulators. Therefore, in order to increase the chances of finding 'yeast sororin' or another previously unidentified cohesin regulatory protein, a more directed screen was attempted in this thesis.

The screen performed in this study was based on the assumption that any yeast sororin candidates (or other novel cohesin regulators) will display a cohesin-like subcellular localisation. It is difficult to determine whether this assumption is justified, as species with sororin orthologues do not exhibit an obvious 'cohesin barrel' for even their core cohesin subunits (Nishiyama et al., 2010). That said, all GFP-tagged budding yeast cohesin subunits form a 'cohesin barrel' recognisable to the trained eye, although its definition is more clear-cut for the core subunits (SMCs, Scc1, Scc3) than for less stably associated, regulatory subunits such as Wapl (Chan et al., 2012). Therefore, any novel regulatory proteins that act on cohesin will presumably at least transiently associate with chromatin-bound rings in the pericentromeric regions, allowing for their detection by live-cell imaging.

Rather than screening all ~5,000 identified yeast genes, the screen performed in this study was streamlined by adding the assumption that any cohesin-

associated protein must spend at least part of its lifetime in the cell co-existing with the intact cohesin complex in order to perform its function. Due to the absence of a defined prophase pathway in yeast, *SCC1* needs to be re-transcribed and re-translated in late G1 following its complete proteolysis at the previous anaphase (Guacci et al., 1997; Michaelis et al., 1997; Uhlmann et al., 1999). Despite the protein levels of other cohesin subunits remaining stable throughout the cell cycle, it was established that all but two (Wapl and Scc2) cohesin subunits have periodically synthesised mRNA transcripts. This was done in collaboration with Prof Patrick Cramer (Ludwig Maximilian University of Munich, Germany) by using dynamic transcriptome analysis (DTA) to measure the steady-state and newly-synthesised mRNA pools during three budding yeast cell cycles after cell synchronisation in G1 (Eser et al., 2014; Miller et al., 2011).

In organisms with a confirmed sororin orthologue, the protein remains persistently associated with chromatin-bound cohesin, so its mRNA should coexist with cohesin subunit mRNAs for a large part of the cell cycle. It was thus assumed that 'yeast sororin' or any other novel cohesin regulators will have cycling mRNAs of a periodicity similar to the core cohesin subunits. That said, mRNAs coding for Wapl and Scc2 are not periodic. These two proteins interact with cohesin only transiently and at particular stages in the cohesin cycle, so any novel cohesin-associated proteins interacting with the ring in this manner may not be detected during the screen.

No novel cohesin-associated proteins could be identified using an imaging-based approach

The complete list of genes with periodically synthesised mRNAs consisted of 479 genes. Each of these genes was assigned an 'mRNA periodicity' value: this was defined as the point in minutes following release from α-factor arrest at which a given gene's mRNA synthesis levels peaked (Eser et al., 2014). Of the 479 genes, 389 were available as viable yeast strains in a commercial GFP protein fusion collection (Invitrogen). Cohesin subunit mRNAs have a periodicity between 13.5 (*SCC4*) and 18 minutes (*SCC1*; Table 1). 193 of the periodic genes with existing GFP protein fusion strains had a 'cohesin-like' mRNA periodicity of between 12 and 19.5 minutes and these were screened first. The remaining 196 candidates with mRNA periodicities ranging from less than a minute to 57 minutes were analysed later in the screen.

Table 1: Cohesin-associated proteins have cycling mRNAs of defined periodicity.

Gene	mRNA periodicity (min)
SCC4	13.5
SMC3	15
ECO1	15
SCC3	16.5
SMC1	16.5
PDS5	18
SCC1	18

As described previously, GFP-labelled cohesin subunits are nuclear and adopt the distinctive 'cohesin barrel' in dividing cells. This can be loosely described as being an asymmetrical nuclear distribution. Therefore, the first stage of the

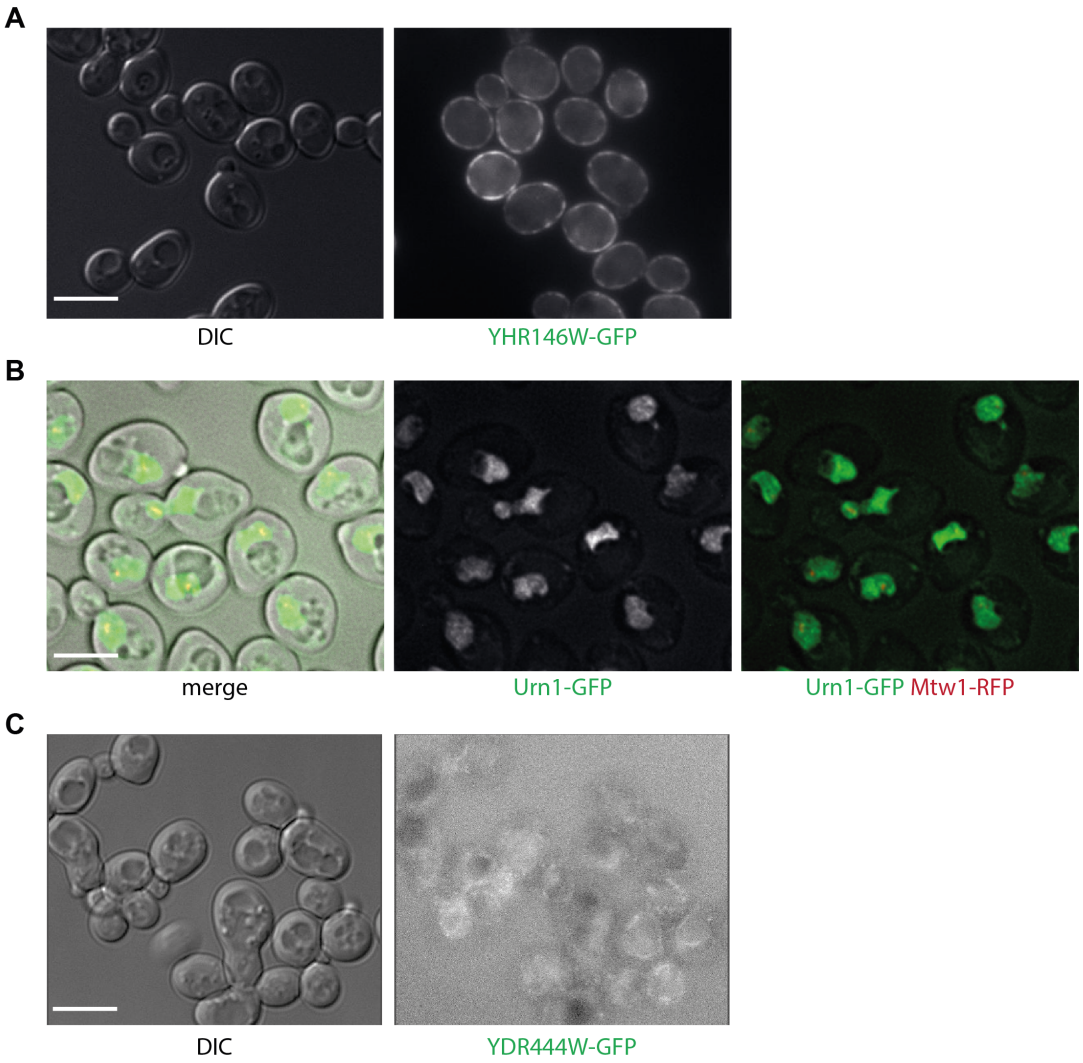

Figure 6. During the primary phase of the screen, candidates were discarded for three main reasons: localisation was non-nuclear or symmetrical nuclear, or the fluorescent signal was too weak.

(A) An example of a discarded non-nuclear protein YHR146W-GFP in a haploid yeast cell.
(B) A discarded 'symmetrical nuclear' protein Urn1-GFP in a homozygous diploid cell with kinetochore marker Mtw1-RFP.
(C) An example of a fluorescent fusion protein YDR444W-GFP with a poor signal-to-noise ratio, making it difficult to unambiguously determine the protein's localisation.

screen aimed to eliminate any non-nuclear proteins (Figure 6A) and those with a symmetrical nuclear distribution (Figure 6B). This was done by visualising live haploid yeast cells in log phase in which the 'cohesin barrel' is not as well-defined as in diploid G2/M cells, but remains recognisable to the trained eye. Each set of genes visualised microscopically during a given session always

contained at least one (blinded) positive control in the form of a strain expressing a GFP-tagged cohesin subunit.

The primary screen of haploid cells excluded the bulk of the 389 genes. Most of the proteins were discarded due to lack of cohesin-like signal distribution but a large proportion of candidates (approximately 20%) had to be abandoned because their fluorescent signal was too weak to accurately assess their subcellular localisation (Figure 6C). For this reason, GFP tagging of the 93 periodic genes absent from the commercially available GFP fusion collection was not attempted, as it was deemed unlikely that functional GFP fusions with good signal-to-noise ratios will be obtained.

The 27 remaining candidates with a clear, asymmetrical nuclear distribution of their GFP signal were then crossed to a strain containing *MTW1-RFP*, which encodes a RFP-labelled kinetochore marker that co-localises with the 'cohesin barrel' (Figure 5A; Chan et al. 2012). These doubly-tagged strains were then crossed with yeast of the same genotype but opposite mating type to create diploid cells homozygous for the GFP-labelled candidate gene and *MTW1-RFP*. Diploid cells grown to log phase were used to enable unambiguous determination of whether the candidates in question indeed localise in a cohesin-like manner.

None of the shortlisted 'yeast sororin' candidates displayed a nuclear distribution matching the 'cohesin barrel' completely, as became evident after lack of co-localisation of any of the 27 shortlisted GFP-fusions with Mtw1-RFP.

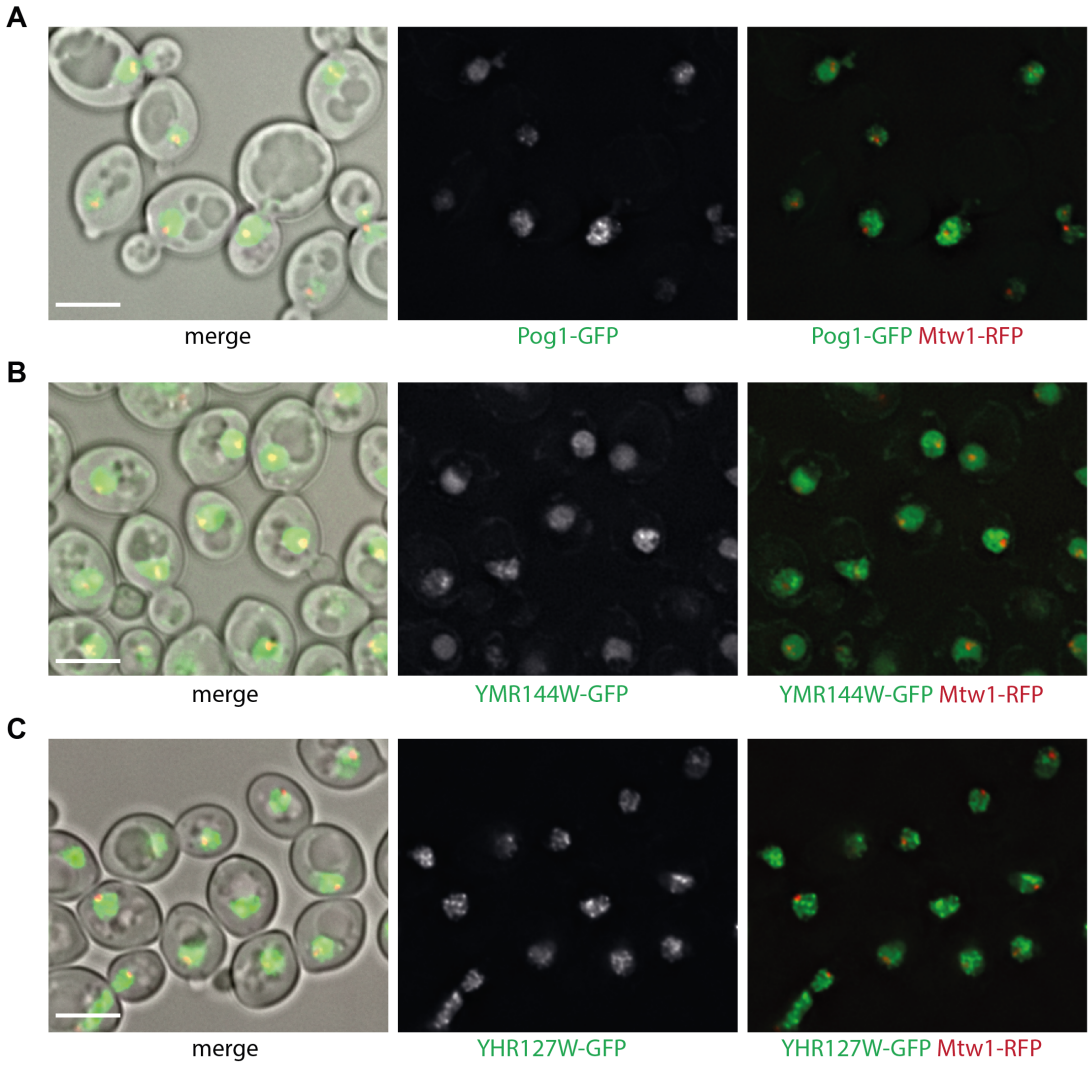

Figure 7. Results of the streamlined GFP fusion screen for 'yeast sororin' did not highlight any convincing candidates for a yeast orthologue of this protein.

No co-localisation of the GFP signal of shortlisted candidate genes with Mtw1-RFP was observed, as exemplified by three proteins that exhibited an asymmetrical nuclear distribution of their GFP signals. Diploid cells homozygous for *MTW1-RFP* and the following GFP fusions were analysed:

(A) *POG1-GFP*
(B) *YMR144W-GFP*
(C) *YHR127W-GFP*

This is illustrated in Figure 7 using three candidates whose asymmetrical nuclear GFP signal distribution was most remnant of the 'cohesin barrel': Pog1, YMR144W, and YHR127W.

Discussion

The negative result of the screen described in this thesis clearly does not prove the absence of sororin in yeast: in this case, "absence of evidence is not evidence of absence" (Sagan and Druyan, 1995). Crucially, there were several key limitations to the experimental approach undertaken in this thesis. Firstly, a more inclusive screen needs to be performed. One cannot assume that sororin will have a periodically expressed mRNA transcript or, if it does, that it was necessarily picked up in the DTA experiment. Clear examples of cohesin regulators without detectably periodic transcripts are Wapl and Scc2. That said, these two proteins interact with the cohesin ring only transiently at precise points in the cell cycle, so their regulation may not be tied to that of cohesin as tightly as it should be for 'yeast sororin,' a protein whose role is to stabilise the ring on chromatin for a prolonged period.

Secondly, any further screening attempted needs to be less reliant on artificially introduced protein modifications. An imaging-based approach is only feasible if all GFP fusion proteins are functional, which will not necessarily be the case: for example, a GFP molecule of 27 kDa is rather large when compared to *X. laevis* sororin of 35 kDa (Rankin et al. 2005). While the GFP-fusion of *X. laevis* sororin behaves like its wild-type counterpart (Nishiyama et al., 2010), addition of a large fluorophore has been shown to interfere with the localisation and function of other similarly-sized proteins (Skube et al., 2010). Even if the GFP-tagged protein is functional, it must generate a signal bright enough for detection and, hence, unambiguous determination of the phenotype. This was not the case for

many of the GFP fusion proteins screened, as many proteins had to be discarded due to insufficient fluorescent signal rather than lack of 'cohesin barrel' *per se*.

Thirdly, any screen performed should be as objective as possible. The nature of imaging is such that it depends on the experimenter observing (and recognising) a particular phenotype. Making the decision of whether a candidate has a symmetric or asymmetric nuclear distribution was not always as trivial as it may seem, especially at low signal levels. What about proteins that were present both in the nucleus and cytoplasm or those that changed localisation dependent on the stage of the cell cycle? Although these considerations were taken into account as best as possible, judgement error will always be part of imaging-based screening.

One idea for a more comprehensive screen for 'yeast sororin' would be a mass spectrometry-based approach. In more complex eukaryotes, the association of sororin with cohesin is dependent on the acetylation of the Smc3 head by Eco1. Thus, one could perform a pull-down on a known cohesin subunit in *waplΔ* and *waplΔ eco1Δ* budding yeast cells. Mass spectrometry could then be used to compare the cohesin-associated proteins in the two genetic backgrounds. The proteome comparison should pinpoint any proteins that are specifically depleted in an *eco1Δ* situation from the cohesin ring (other than Eco1 itself, naturally). This would produce a 'clean' data set, but the purification procedure may distort or even abolish the binding of any novel candidate protein(s).

Another idea is performing in budding yeast a modified version of the original *X. laevis* screen that identified sororin: a screen for proteins whose levels are controlled by APC/C^{Cdh1} (Rankin et al., 2005). In vertebrates, sororin levels are kept low in G1 after its ubiquitin-mediated degradation at the end of the previous cell cycle. Screening for additional targets of APC/C^{Cdh1} should identify proteins degraded at the anaphase to G1 transition, as opposed to APC/C^{Cdc20} that targets proteins like securin at the metaphase to anaphase switch (Rankin, 2005). One way to do this would be to compare the proteomes of wild-type (WT) and *cdh1Δ* yeast cells by mass spectrometry as mentioned previously. Any uncharacterised proteins stabilised by the absence of APC/C^{Cdh1} could be viable candidates for 'yeast sororin.'

Despite the mentioned limitations of the screen performed in this thesis, one must also consider the not unrealistic possibility that there simply is no sororin orthologue in yeast. While the role of vertebrate sororin is to keep cohesin protected against releasing activity, equivalent 'locking' changes to the cohesin ring could be induced by another mechanism in yeast (Nasmyth, 2011). Perhaps how Smc3 acetylation *per se* affects cohesin is sufficient to prevent Wapl's association with and/or action on yeast cohesin. This possibility is appealing, since the releasing complex does not play a central role in *S. cerevisiae* as it does in more complex eukaryotes (there is no defined prophase pathway in yeast), so neutralising the releasing reaction may require less dedicated cellular machinery. An extension of this hypothesis would be the proposition that there is something fundamentally different about the mechanism of cohesion in *S. cerevisiae*. One way to investigate this would be to set up an *in vitro* assay

capable of answering questions about molecular details of cohesion difficult to address *in vivo*. The aim of the next section of this thesis was to purify a high-quality recombinant yeast cohesin complex for use in biochemical assays.

CHAPTER III

Biochemical reconstitution of the *S. cerevisiae* cohesin complex

CHAPTER III: Biochemical reconstitution of the *S. cerevisiae* cohesin complex

---

Introduction

If *S. cerevisiae* indeed has no sororin orthologue, then how does it robustly establish cohesion time and time again? An explanation for this would be that there is something fundamentally different about the mechanism of cohesion between budding yeast and more complex eukaryotes. As mentioned before, one of the most elegant ways to address this would be to create an *in vitro* assay that could be used to establish the key players and minimum set of components required for each step of sister chromatid cohesion – cohesin loading, cohesion establishment, and the releasing reaction. Once functional, an *in vitro* assay for cohesion should also be capable of addressing questions that are difficult to unequivocally answer *in vivo*. For example, is the cohesin hinge the DNA 'entry gate'? Is Smc3 acetylation itself sufficient for cohesion establishment? Is ATP required for the cleavage-independent release of cohesin from chromatin? It should be then possible to determine whether there are any major differences in the mechanism of cohesion in various organisms.

An *in vitro* assay for the cohesin loading reaction in *S. pombe* has recently been used to show that Scc2 interacts with the cohesin ring at multiple points and

promotes efficient loading by stimulating cohesin's ATPase activity (Murayama and Uhlmann, 2014). While valuable and informative, the published *S. pombe* assay cannot give direct insight into the molecular details of cohesin's function in *S. cerevisiae* because of fundamental differences in the mechanism of cohesion between the two yeasts. In *S. pombe*, the two lysine residues on Smc3 acetylated by Eco1 to counteract Wapl activity are not essential and neither is Pds5 (Feytout et al., 2011; Wang et al., 2002). Establishing a biochemical assay using budding yeast components could therefore address such discrepancies, especially given that there is a broad range of tools available for investigation of cohesion using this model system (e.g. site-specific cross-linkable residues in each of cohesin's interfaces).

The aim of this study was to set up a reliable and flexible purification method for the yeast core cohesin tetramer for future use in biochemical assays. The tetrameric *S. pombe* cohesin complex used in the loading assay mentioned above was obtained by overexpression of gene pairs integrated into the *S. cerevisiae* genome (Murayama and Uhlmann, 2014). This purification method was not adopted here, because the information had not yet been published when work on this project was initiated. Even if it were, the purification strategy adopted in this thesis should arguably give the system more flexibility and be more time-effective.

The most notable biochemical assay for which the yeast cohesin tetramer isolation was attempted is an *in vitro* assay for cohesin release, which could be used to determine the minimum requirements for the releasing reaction.

Because the final goal of a functional biochemical assay lies at the end of a long series of steps, the reconstituted recombinant cohesin complex could also be used in multiple investigations that aim to further our understanding of cohesion:

(i) *Visualising cohesin by cryo-electron microscopy (cryo-EM)*. Cryo-EM is a technique that enables visualisation of large protein complexes in their native environment with great atomic detail: resolution can reach below 1 nm in some cases (Hoenger, 2014; Schur et al., 2013). So far, 'traditional' EM had been used to visualise cohesin holocomplexes but little is known about the conformation of Scc1 and Scc3 when part of the cohesin complex (Anderson et al., 2002; Melby et al., 1998). Cryo-EM could be used to visualise the cohesin tetramer as a whole, or its subcomplexes, in greater molecular detail and increase our understanding of the complex's architecture. One could investigate, for example, whether binding of Scc3 to the cohesin trimer influences its conformation.

(ii) *Chemical cross-linking to identify the role of each cohesin interface*. While available evidence gives us a good idea about the identity of cohesin's DNA 'entry' and 'exit gates', there is still no direct proof due to limitations of the techniques used. A biochemical assay with purified cohesin components is perhaps the only way to definitively show that the ring topologically entraps DNA and which interfaces must open for its loading and release. To identify cohesin's DNA 'entry gate', each interface in turn should be chemically cross-linked prior to performing the *in vitro* loading reaction. The version of cohesin unable to load would demonstrate which interface needs to open to allow DNA

entry into the ring. To confirm the topological embrace model, cohesin/DNA complexes should be recovered by immunoprecipitation, all three interfaces should be cross-linked after cohesin loading, all proteins denatured, and cohesin/DNA topology determined by agarose gel electrophoresis. Lastly, once the releasing assay is functional, cross-linking each interface and determining which version prevents cohesin release after addition of purified releasing complex subunits would reveal the identity of the DNA 'exit gate.'

(iii) *Using Förster Resonance Energy Transfer (FRET) approaches to monitor the dynamics of cohesin's interfaces*. Single-molecule FRET-based approaches could be used to monitor opening and closing of cohesin's interfaces during *in vitro* loading and release. This would require introduction of different fluorescent dye pairs at each interface and monitoring of distance-dependent energy transfer between fluorophores. It should be theoretically possible to detect opening and closing of cohesin's interfaces in real-time, which would provide invaluable insight into the dynamics of cohesin's interaction with chromatin.

Strategies for reconstituting the *S. cerevisiae* cohesin complex

The core cohesin complex is composed of four protein subunits: in *S. cerevisiae*, these are Smc1, Smc3, Scc1, and Scc3. Of these, only yeast Scc3 has previously been reliably purified as a full-length protein using a bacterial expression system (Roig et al., 2014). *E. coli* has been previously used to express and isolate recombinant SMC hinges and short fragments for crystallisation, while

purification of full-length yeast SMCs has been described using an insect cell expression system (Gligoris et al., 2014; Haering et al., 2002; Kurze et al., 2011). The full-length SMCs co-purified robustly and material obtained was of high enough yield and quality for electron microscopic analysis. These findings indicated that full-length versions of Smc1 and Smc3 are 'well-behaved' during biochemical isolation, which is presumably at least partially due to their well-defined architecture and distinct functional domains, indicating that SMC dimer co-purification may be a viable option.

Previous attempts in the lab to purify budding yeast Scc1 revealed that the protein is largely insoluble and any soluble material is susceptible to non-specific proteolytic cleavage. Bioinformatic approaches have identified few structured domains in Scc1's protein sequence. In fact, only short stretches of Scc1's sequence (one in the N terminus and two towards the C terminus) are predicted to assume alpha-helical structures, in addition to one short beta-strands at the N terminus and two at the very C terminus of the protein (Jones, 1999). In other words, the bulk of the protein is free from any secondary structure motifs recognisable by bioinformatic algorithms. These predictions are largely supported by conservation analysis, where the most evolutionarily conserved residues between yeasts and more complex eukaryotes are located at the N terminus of Scc1 and, to a lesser extent, at its C terminus (Figure 8; Larkin et al. 2007). In fact, these data also tie in well with what is known about Scc1's interaction with its binding partners Smc1 and Smc3 (binding via Scc1's C and N termini, respectively) and its conserved role during anaphase (cleavage by

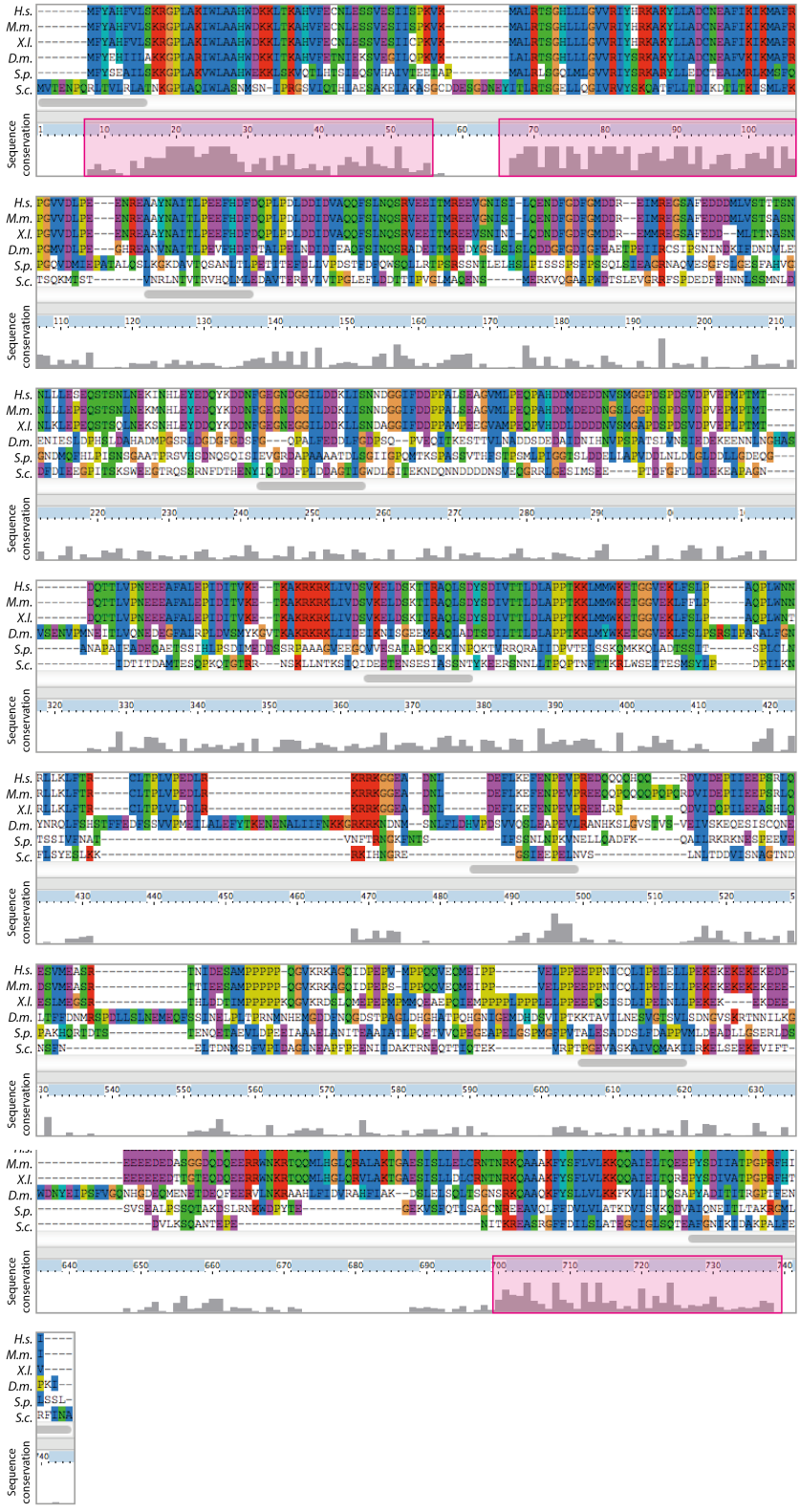

Figure 8. Alignment of Scc1 between yeasts and more complex eukaryotes revealed areas of conservation at the very N and C termini of the protein.

Areas of highest sequence conservation have been highlighted in magenta boxes.
[UniProt identifiers are given in brackets: *H.s.* – *Homo sapiens* (O60216); *M.m.* – *Mus musculus* (Q61550); *X.l.* – *Xenopus laevis* (O93310); *D.m.* – *Drosophila melanogaster* (Q9U6D9); *S.p.* – *Schizosaccharomyces pombe* (P30776); *S.c.* – *Saccharomyces cerevisiae* (Q12158).]

separase within the flexible loop between Scc1's two termini). Combined, these observations led to the idea that a major fragment of Scc1 is unstructured and flexible in solution.

It is not unfeasible to assume that part of the reason Scc1 has not been successfully purified so far is a direct consequence of the large, unstructured region of the protein. Firstly, the low solubility of recombinant Scc1 can be partially explained by its unfavourable thermodynamic interactions: in a structured, globular domain, hydrophobic residues are excluded from the aqueous environment by hydrophilic residues clustered on the outside, but this is not the case in unstructured domains, where hydrophobic residues often remain exposed. When extreme, exclusion of non-polar amino acids by water molecules can cause the protein to precipitate out of solution and, hence, become unusable in biochemical assays. Secondly, suboptimal polar/non-polar interactions can also promote protein misfolding (Song, 2013): it is believed that the main factors promoting correct folding of cytoplasmic proteins are the solvent-excluding clustering interactions between a protein's non-polar side chains (Chandler, 2005). While this effect is largely offset *in vivo* by molecular chaperones that assist protein folding (Saibil, 2013), introducing an exogenous, eukaryotic protein into a prokaryotic host creates a scenario where a protein only folds correctly if it can do so either spontaneously or using bacterial chaperones.

In order to maximise the chances of obtaining a functional cohesin complex within the timeframe of this thesis, several strategies were attempted:

1. Co-expression of Scc1 with Smc1 prior to supplementation with the Smc3 monomer (and later Scc3).
2. Expression of Scc1 alone followed by addition of the Smc1/Smc3 heterodimer (and later Scc3).
3. Co-expression of Scc1 with Scc3 for supplementation with the Smc1/Smc3 heterodimer (co-expressed separately to Scc1 and Scc3).

Functional complex assembly was to be verified by one of two methods: co-immunoprecipitation (co-IP) and site-specific chemical cross-linking. So far, the interaction of Scc3 with the cohesin trimer has not been captured by chemical cross-linking, although it is known this binding occurs via Scc1 (Roig et al., 2014). A co-IP experiment is therefore the only viable option to determine whether Scc3 associates specifically with the recombinant complex. Naturally, the pull-down should be attempted reciprocally, i.e. by pulling on Scc3 itself as well as on one of the components of the cohesin trimer, to see if all four components can be recovered. Data obtained from these co-IPs should not only determine whether Scc3 interacts with the trimer but will also provide a good indication of whether the Smc1/Smc3/Scc1 trimer itself associates correctly.

In the case of the trimer, functional complex assembly should be further verified by chemically cross-linking of the cohesin ring by introducing cross-linkable cysteine pairs at each of the three cohesin interfaces: Smc1/Smc3, Smc1/Scc1, and Smc3/Scc1. These cysteine pairs were designed based on the crystal structures of the three interfaces and have been used previously to address cohesin topology (Bürmann et al., 2013; Farcas et al., 2011; Haering et al., 2008;

Kurze et al., 2011). Since chemical cross-linking can only occur within a narrow range of distances, cross-linked cysteine bridges formed between subunits will give a good indication that the cohesin complex is in its wild-type conformation and, hence, likely to be functional.

'Bi-cistronic' expression of codon-optimised *S. cerevisiae* proteins in *E. coli*

The BL21 DE3 *E. coli* strain was used as the preferred protein expression system for three main reasons. Firstly, it is a well-established protein purification system with readily available tools and, if a method needs troubleshooting, a vast amount of expertise is available. Secondly, the system allows great flexibility in terms of construct design, including a wide choice of inducible expression vectors and affinity tags. Thirdly, the period of time between construct design and commencing protein purification is relatively short compared to other expression systems.

All but two (tryptophan and methionine) naturally occurring amino acids are coded for by more than one codon (Ikemura, 1985). Theoretically, synonymous codons for a given amino acid should occur at equal frequencies in an organism's genome, since translation of any of them would result in the same protein product. This is not the case, however, as many organisms exhibit a 'preference' towards a particular synonymous codon in a process termed codon usage bias. In fast-growing microbes especially, codon bias matches the tRNA pool available in

the cell, suggesting that this phenomenon may have evolved to increase translation yield and accuracy (Grantham et al., 1980; Ikemura, 1981; Plotkin and Kudla, 2011). While its precise evolutionary roles are still under debate, codon usage bias has one corollary that is key to this study: protein expression can be increased many-fold through use of synonymous codons favoured by the expression host (Gustafsson et al., 2004; Plotkin and Kudla, 2011).

Since a method for reliable expression of non-codon-optimised yeast *SCC3* in *E. coli* has already been established in the lab, the purification of Scc3 was not the main concern of this study (Roig et al., 2014). Rather, this thesis focussed on purification of the cohesin trimer (Smc1, Smc3, and Scc1) with addition of Scc3 at later stages to complete the core cohesin tetramer. Thus, the three *S. cerevisiae* cohesin genes *SMC1*, *SMC3*, and *SCC1* were codon-optimised for *E. coli* (GenScript). All constructs were then expressed from the pET28a vector in BL21 DE3 *E. coli* cells. Expression of genes cloned into this vector was driven by the IPTG-inducible T7lac promoter during a slow induction (15-18 hours) at 16°C. Low growth temperature during protein overexpression was used for three reasons: (i) to promote correct folding by increasing the amount of bacterial chaperones present, (ii) to increase the solubility by minimising inclusion body formation, and (iii) to limit degradation by reducing the activity of heat shock proteases (Ferrer et al., 2003; Sahdev et al., 2008; San-Miguel et al., 2013; Vera et al., 2007).

For co-expression of two proteins, the two coding sequences were arranged in series in a single pET28a plasmid to generate a 'bi-cistronic' construct, whereby

two proteins are expressed individually from a single mRNA transcript controlled by a single promoter. To ensure the latter of the two genes was translated as well as the first, a short ribosome binding sequence (RBS) was added at the 5' end of the second gene. This approach both avoided the issue of incompatibility when using multiple vectors containing the same origin of replication (Velappan et al., 2007) and attempted to preserve a stoichiometric ratio between any two components expressed. Furthermore, purifying two proteins simultaneously facilitated labelling of individual subunits without overcomplicating the system with many expression vectors. This becomes important for approaches such as single-molecule FRET, where introduction and testing of multiple fluorescent dye pairs at each cohesin interface is required. This would be achieved through insertion of unnatural amino acids at various position in the cohesin ring prior to chemically attaching the fluorophore, a process that should be greatly accelerated by a simple cloning and expression set-up (Singh-Blom et al., 2013).

The next step in designing the expression constructs was to determine which affinity tags to use. Ideally, each of the tetramer's component proteins should be fused to a unique affinity tag such that each subunit can be visualised individually by Western blotting when complex assembly is verified by co-IP, as mentioned previously. However, this approach would also mean that four separate purification protocols would need to be optimised. This was determined to be impractical, so a compromise was made to use a maximum of three unique affinity tags. It is important to note that size difference between the four expressed proteins is large enough to detect them all as individual bands on

a Western blot even when using only three distinctive tags (each SMC is 142 kDa, Scc3 is 134 kDa, and Scc1 is 64 kDa).

The range of affinity tags available for reliable purification of recombinant proteins from bacterial expression systems is vast (Terpe, 2003). When choosing an affinity tag for a particular application, one must consider several parameters including the tag's ability to enhance a protein's solubility (if required), the tag's length and specificity, the affinity resin's optimal binding and elution conditions, and the cost of the resin (Waugh, 2005). Since functional cohesin trimer formation relies on interactions between the N and C termini of each Smc1, Smc3, and Scc1, the most important consideration in their purification was therefore deemed to be affinity tag size. Short tags are less likely to interfere with target protein folding and, hence, function.

The StrepII, six-histidine (6×His), and FLAG tags were therefore chosen due to their small size (8, 6, and 6 amino acids, respectively) and readily available purification procedures. Additionally, purification using a StrepII and 6×His tag can be scaled up with relative ease if needed; the high cost of the FLAG affinity resin and eluting FLAG peptide makes this resin better suited to smaller-scale purifications. The StrepII tag was a particularly attractive choice, since the affinity resin is highly specific and can also tolerate the presence of EDTA, a chelating agent that works to inhibit proteases and thus protects full-length target products. The binding buffers between each of the three chosen resins are largely compatible, so a tandem purification approach remained viable if required.

Recombinant yeast Scc1 can be co-purified with Smc1 but Smc3 cannot be purified as a monomer to complete the cohesin trimer

Scc1's C terminus, but not its N terminus, has been shown to have the ability to interact with the Smc1/Smc3 heterodimer *de novo*: the association of Scc1's C terminus with the Smc1 head can occur independently of the binding status of Scc1's N terminus to Smc3. Indeed, the association of Scc1's N terminus with Smc3 depends on its C terminus being already in contact with an Smc1 head bound by ATP (Arumugam et al., 2003). Following this logic, co-expressing Scc1 with Smc1 should generate a stable dimer, poised for association with Smc3, which would complete the trimeric cohesin ring.

The Scc1/Smc1 expression vector was constructed in a pET28a vector with Scc1-6×His being followed by Smc1-FLAG in the plasmid, with a RBS before the second gene in the sequence (Figure 9A). Small-scale induction tests of BL21 DE3 *E. coli* cells containing this construct demonstrated that both proteins could be detected in the whole-cell extract (WCE) by Western blotting. Despite being codon-optimised, neither Scc1-6×His nor Smc1-FLAG were expressed at levels high enough to easily distinguish them in a Coomassie-stained gel (Figure 9B).

3.5 litres of cells expressing Scc1-6×His and Smc1-FLAG were harvested and a His-tag purification was performed using 5 ml of affinity resin. Proteins were competitively eluted off the resin using a continuous gradient of increasing imidazole concentration (0 mM to 500 mM) spread across 40 ml with an

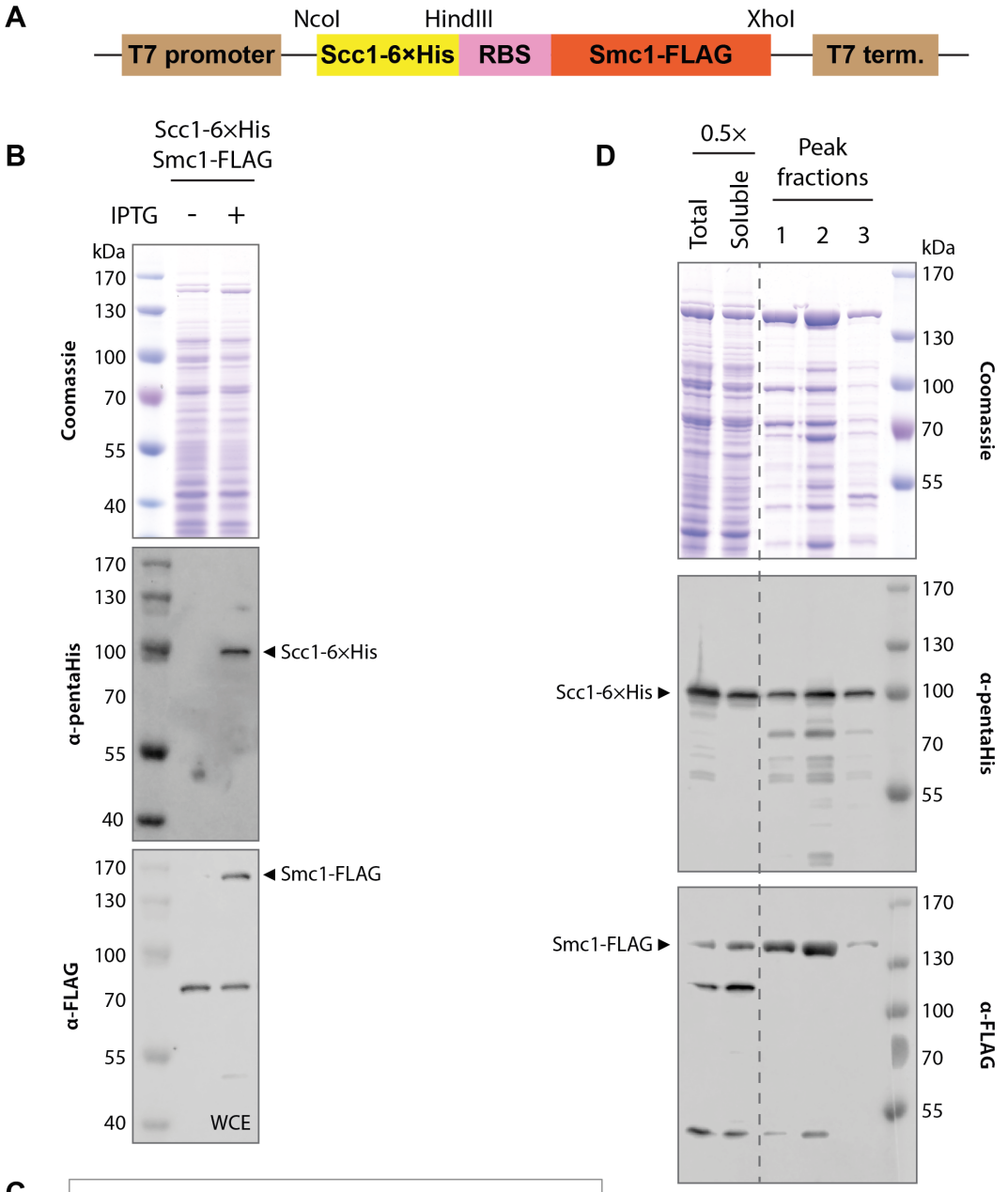

Figure 9. Recombinant yeast Scc1-6×His can be co-purified with Smc1-FLAG after codon-optimisation of the genes in question for the bacterial expression host.

(A) Schematic representation of the Scc1/Smc1 construct used for protein expression (RBS: ribosome binding site, circa 30 nucleotides).
(B) IPTG induction of *E. coli* cells hosting the expression vector confirmed the presence of full-length versions of both Scc1 and Smc1 to levels detectable in the whole cell extract (WCE) by Western blotting but less apparent in a Coomassie-stained gel.
(C) Elution of material bound from 3.5 litres of starting culture from a His-tag affinity column produced a large, sharp absorbance peak at 280 nm (A280).
(D) Appropriate elution fractions (marked in (C) with blue rings) were analysed by SDS-PAGE followed Western blotting to confirm the presence of both full-length peptides.

additional 25 ml wash at the highest imidazole concentration. A single, sharp peak of material absorbing at 280 nm was observed (Figure 9C). Fractions from this peak contained both Scc1-6×His and Smc1-FLAG, as detected by Western blotting, but Coomassie staining revealed that the elution also contained many protein impurities (Figure 9D).

Judging by the ladder of signal visible in anti-His Western blot, some of the extra protein bands detected could be attributed to N terminal degradation products of Scc1-6×His. That said, certain bands did not correspond to any Western blot signals and therefore were assumed to be either products resulting from C terminal cleavage that removed the epitopes from the recombinant proteins or contaminating proteins from the host strain. One band in particular (the strongly staining band running just below the 70 kDa marker) appeared in almost every His-tag affinity purification. In collaboration with Dr Benjamin Thomas (Central Proteomics Facility, Sir William Dunn School of Pathology, University of Oxford), mass spectrometry was used to reveal that the band in question was mainly composed of glutamine-fructose-6-phosphate aminotransferase, a bacterial enzyme with histidine-rich patches clearly capable of interacting with the affinity resin used in this study.

Before proceeding to another round of purification of the Scc1/Smc1 eluate, it was imperative to determine whether the remaining component of the trimer, Smc3, can be purified on its own. While Smc3 is predicted not to form homodimers, as its hinge domain interacts exclusively with the Smc1 hinge region, purification of monomeric, recombinant Smc3 only been reported in the

literature using an insect cell expression system (Haering et al., 2002). An Smc3-StrepII construct in pET28a was created (Figure 10A) and its expression verified by induction of a small culture (Figure 10B). Similarly to the Scc1/Smc1 construct, Smc3-StrepII induction was clearly visible on a Western blot, but less so on a Coomassie stained gel of the WCE. 5 ml of StrepII resin was then used to affinity purify protein from 1.6 litres of induced culture. Material was eluted off the column using a continuous gradient of increasing d-Desthiobiotin concentration (0 mM to 10 mM) spread across 40 ml with an extra 25 ml wash at the top end of the gradient.

Two small absorbance peaks and one equally small, wide peak were observed at the very beginning of the elution gradient (Figure 10C). Western blotting confirmed that Smc3-StrepII was present in all three fractions, but the yield was very poor at below 0.1 mg/ml. What is more, there was very little depletion of soluble material, as can be seen when comparing the 'soluble' and flow-through 'FT' lanes (Figure 10D). One explanation for this is the use of pre-packed affinity resin cartridges as opposed to batch treatment. The cell lysate was passed through the column at a relatively fast flow rate of 0.5 ml/min: this treatment reduces the amount of non-specific binding to the resin, resulting in a cleaner elution. The flipside is that specific binding is also reduced, characterised by detection of material in the FT. A solution would be to run the lysate multiple times through the same column until there was marked depletion of soluble material.

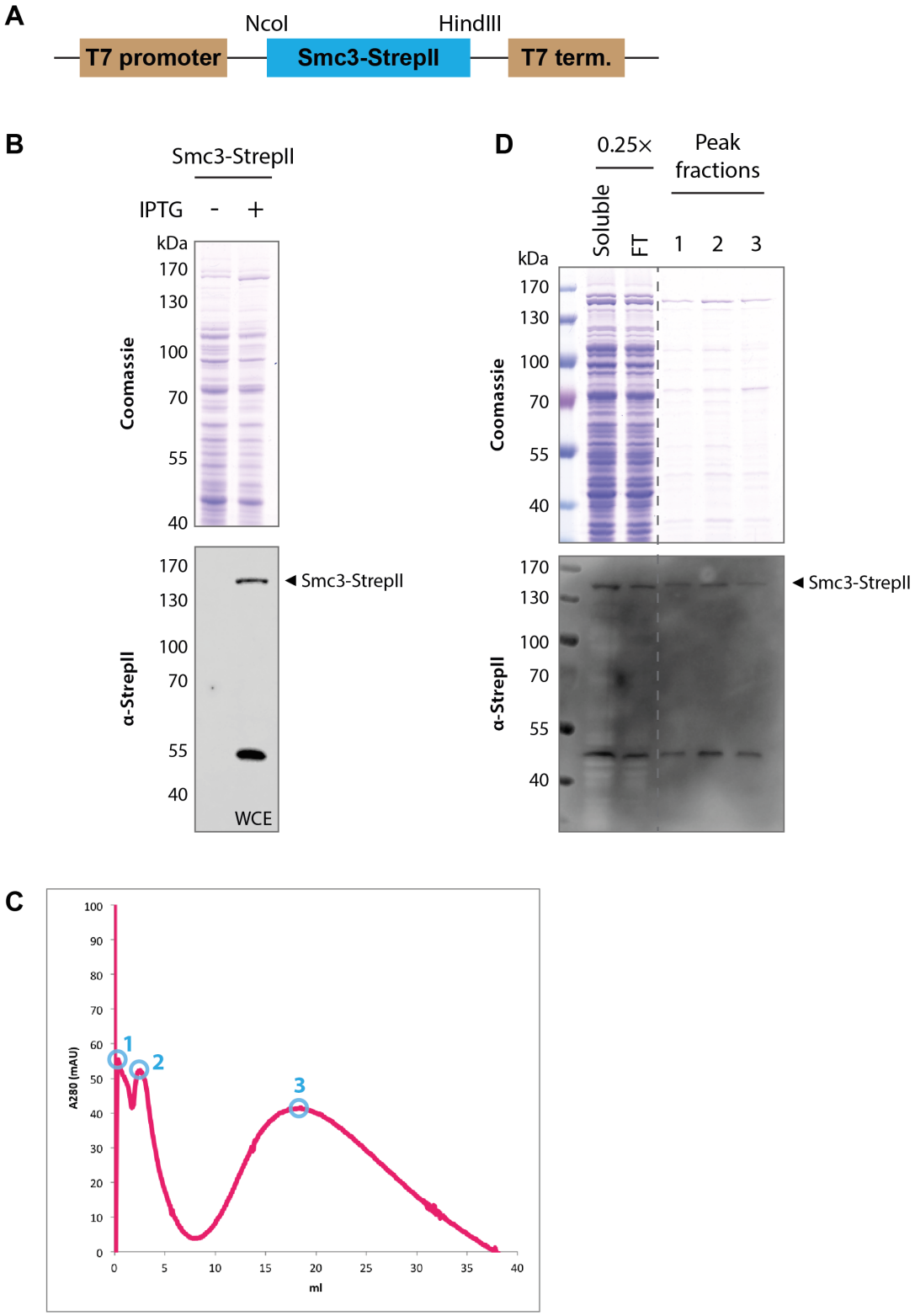

Figure 10. Monomeric Smc3-StrepII cannot be reliably purified using the bacterial expression system used in this thesis.

(A) Schematic representation of the Smc3 construct introduced into the pET28a expression vector.
(B) Western blotting confirmed the presence of full-length Smc3 in the expression strain's WCE following induction, but its levels were too low to easily detect in a Coomassie-stained gel.
(C) Despite 1.6 litres of starting material, the absorbance peak generated during elution from a StrepII-tag affinity column was small and broad.
(D) Three elution fractions (marked in (C) with blue rings) were analysed by SDS-PAGE followed Western blotting. Smc3-StrepII could be detected in all fractions in minute amounts.

Successful purification of the Smc3 monomer was critical to reconstituting the yeast cohesin trimer using the previously co-purified Smc1 and Scc1. There could be a number of reasons as to why the Smc3 protein expressed was not giving a single, sharp absorbance peak with good yields. These could range from a technical issue such as adsorption of protein to the column filter and hence reduction in recovered material, to a more a inherent problem such as protein multimerisation or aggregation obscuring the affinity tag from binding. Given these concerns, it was decided to pursue a more promising strategy of reconstituting the cohesin complex.

Single-step purification of the SMC heterodimer reveals that Smc1 can be reciprocally co-purified with Smc3

Smc1 and Smc3 have been shown to heterodimerise through interactions between their respective 'hinge' domains. This has previously enabled the co-purification of the two full-length *S. cerevisiae* proteins from insect cells and their recombinant hinge fragments from bacterial expression systems (Haering et al., 2002; Kurze et al., 2011). The robust interaction between the Smc1/Smc3 hinges led to the assumption that, if their expression is high enough and non-specific proteolysis is minimised, recombinant versions of the two yeast proteins should co-purify even when expressed in a bacterial host cell. This heterodimer would then be supplemented with Scc1 and Scc3 and the functional interaction of the two subcomplexes verified to reconstitute the cohesin tetramer.

As before, a 'bi-cistronic' construct was created for the SMC heterodimer: Smc3-StrepII followed by a RBS and Smc1-6×His (Figure 11A). It was not certain whether the two proteins were going to be expressed at equal levels, so this arrangement of affinity tags enabled a tandem purification to be performed if required. The construct was introduced into the *E. coli* expression strain and induced at a low temperature. Western blotting confirmed the presence of both proteins in the WCE of the host strain, with maximal expression being achieved after an overnight incubation with IPTG as opposed to a short two-hour induction (Figure 11B). It was also noted that Smc1-6×His, the second protein in the construct, was expressed at high levels even in the absence of IPTG. While this phenomenon did not occur with the Scc1-6×His/Smc1-FLAG vector, it was not detrimental to the SMC heterodimer experiment so was attributed to a random mutation in the pET28a backbone.

In anticipation of a non-stoichiometric ratio of the two SMCs, both a His-tag (Smc1) and StrepII-tag (Smc3) affinity purification was attempted separately. The His-tag pull-down on Smc1 revealed a large, sharp absorbance peak originating from 2 litres of starting culture (Figure 11C) and it was confirmed to contain both Smc1 and Smc3 by SDS-PAGE and Western blotting (Figure 11D). As with the Scc1/Smc1 His-tag purification, there was a substantial amount of contaminants in the eluate, much of which was presumably bacterial in origin, but could also have been partially composed of products generated by non-specific proteolysis.

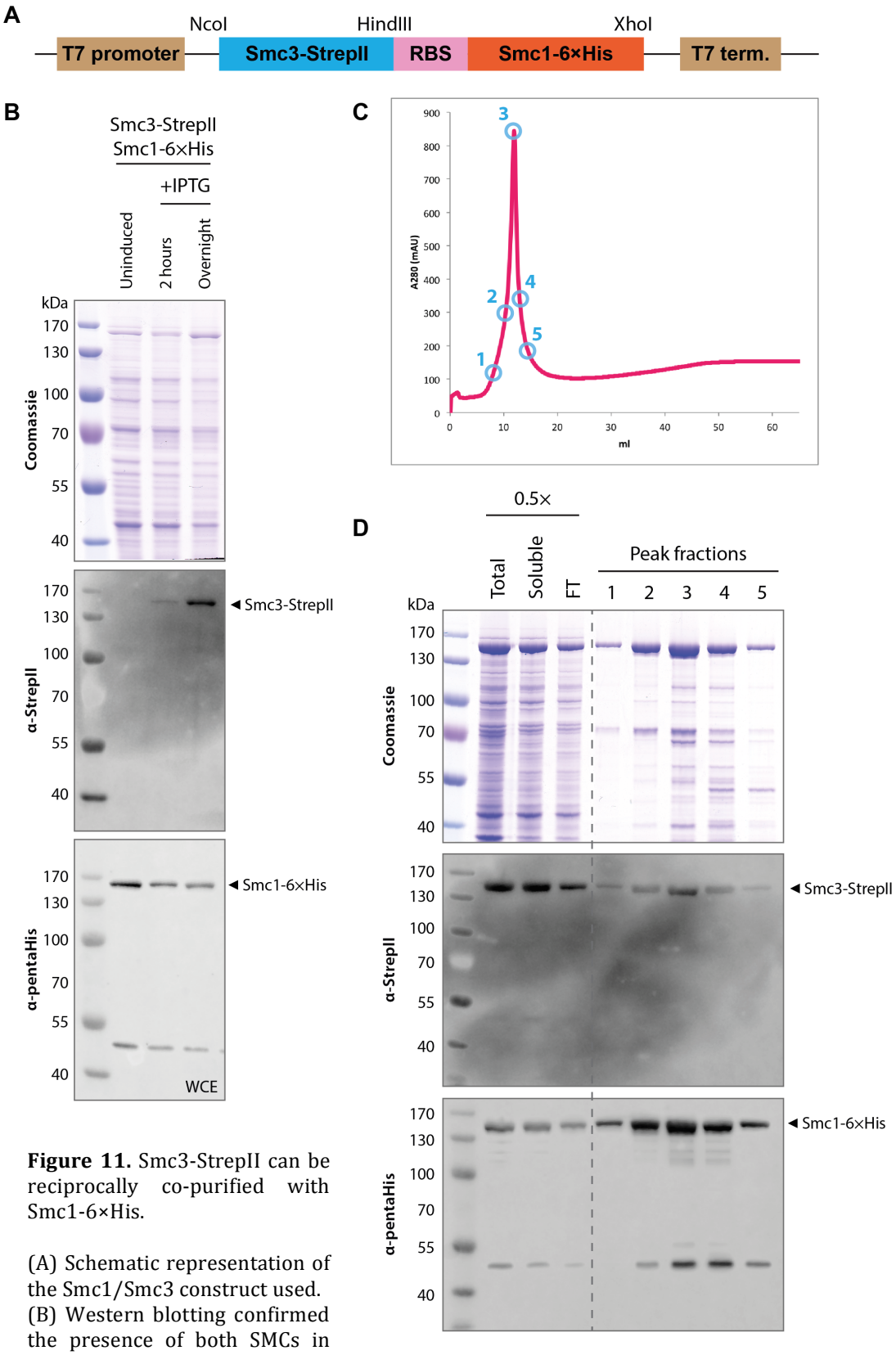

Figure 11. Smc3-StrepII can be reciprocally co-purified with Smc1-6×His.

(A) Schematic representation of the Smc1/Smc3 construct used.
(B) Western blotting confirmed the presence of both SMCs in the WCE. Presumably due to a mutation in the expression vector, Smc1 was rendered unresponsive to the induction and remains at high levels even in the absence of IPTG.
(C) His-tag affinity purification from 2 litres of induced culture generated a sharp absorbance peak.
(D) Fractions marked in (C) with blue circles contained both Smc1-6×His and Smc3-StrepII, as determined by Western blotting.

[Figure 11 continued on next page.]

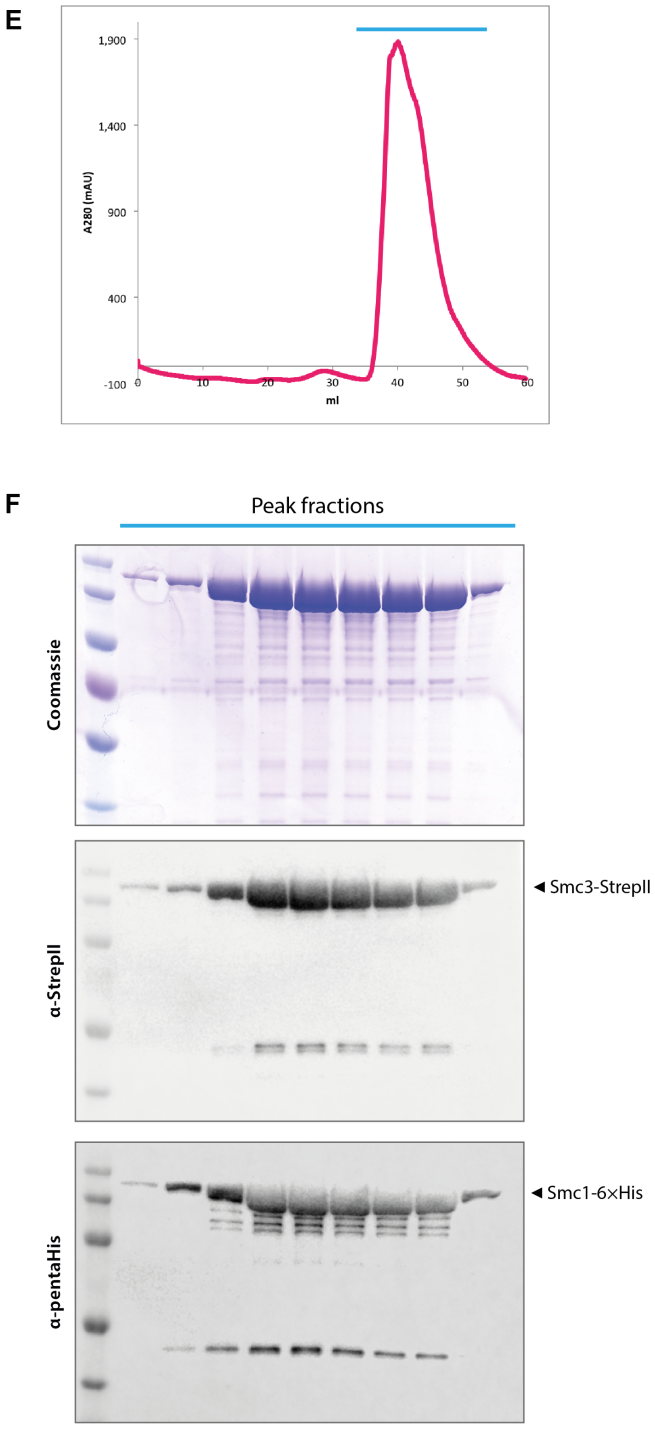

Figure 11. Smc3-StrepII can be reciprocally co-purified with Smc1-6×His.

(E) StrepII-tag affinity purification using 3 litres of bacterial culture produced a large peak, broader than that observed in (C). Fractions for analysis were chosen across the peak marked by the blue line. [Note: the StrepII-tag peak appeared much later during the gradient elution because a pump wash in elution buffer had been omitted.]
(F) The peak in (E) contained a large amount of both Smc1 and Smc3, with a relatively low background of contaminating proteins.

The reciprocal, StrepII-tag pull-down on Smc3 using 3 litres of induced cells resulted in a large, broad absorbance peak (Figure 11E). The peak's broadness may been a consequence of two peaks not becoming fully resolved. That said, all fractions throughout the peak contai ned a large (and apparently homogeneous) amount of both SMCs (Figure 11F). The StrepII-tag was cleaner than the His-tag elution: the Coomassie-stained gels demonstrate that the ratio between the proteins of interest and contaminating background bands is lower in the His-tag purification. This was partially due to the higher specificity of the StrepII-tag, but presumably also due to the presence of EDTA in the StrepII-tag lysis buffer, which limited any non-specific proteolysis.

The next point to address was whether either purification method of the Smc1/Smc3 dimer resulted in the formation of a stoichiometric heterodimer that is hence likely to be functional. There are several widely-accepted methods to approach this:

(i) *Coomassie staining or Western blotting against the same epitope*. If the proteins in question can be distinguished on the basis of their size, their visualisation within a polyacrylamide gel or by Western blotting against the same epitope can give a good estimate of their relative abundances. However, both yeast SMCs have a molecular mass (MW) of 142 kDa, so this was clearly not an option.

(ii) *Mass spectrometry-based approaches*. Quantitative mass spectrometry can be used to label complexes with stable isotopes to determine their relative amounts

in a sample. This is a timely and costly approach, but methods of label-free stoichiometrical analysis have also been developed recently (Smits et al., 2013).

(iii) *Multi-Angle Light Scattering (MALS)*. The stoichiometry between the protein components within a complex can be calculated from accurate measurements of its molecular mass following size exclusion chromatography (Mogridge, 2004).

Because of the abundance of data in the literature regarding the SMC proteins, a more directed and, ultimately, more informative experimental approach was undertaken to assess the quality of the purified Smc1/Smc3 material. This method consisted of introducing a specific cysteine pair in the Smc1/Smc3 hinge prior to attempting to cross-link the interface using bi-functional thiol-reactive chemical cross-linkers. Since chemical cross-linking can only occur across very short distances, the appearance of a cross-linked product would imply that both SMC hinges are functionally assembled and interacting physiologically. Hence, at least part of the eluate would have to be forming a heterodimeric complex with an Smc1 : Smc3 ratio of 1:1.

Site-specific cross-linking of the SMC hinge interface demonstrates that both purified SMC heterodimers are functionally assembled

The cysteine pair introduced into the SMC hinge had been designed based on the interface's crystal structure and had been previously used to address its topology (Kurze et al. 2011). Site-directed mutagenesis (SDM) was used to make a single

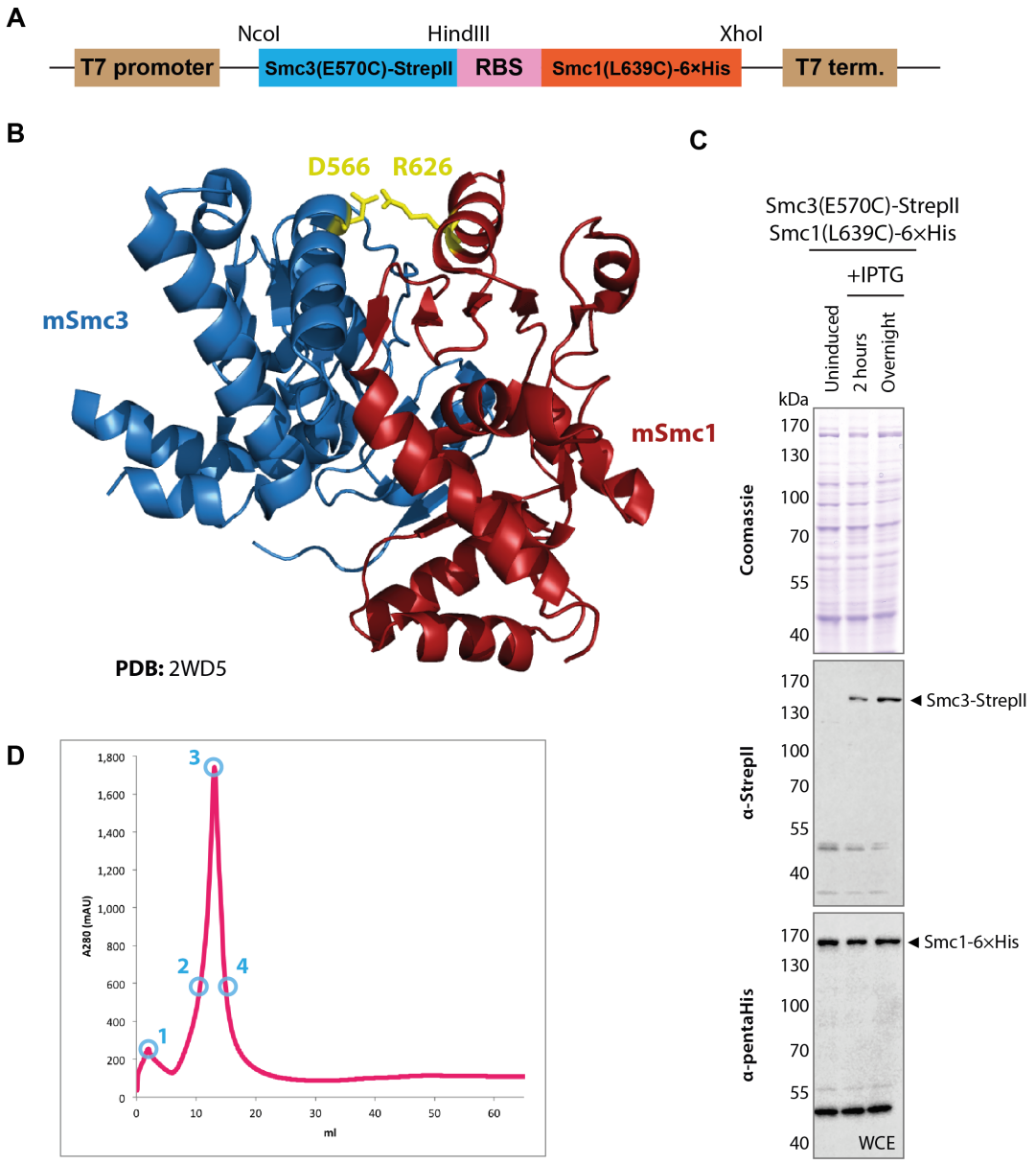

Figure 12. Substitution of E570 in Smc3 and L639 in Smc1 with cysteine residues does not affect the two proteins' co-purification.

(A) Diagram of the Smc1(L639C)/Smc3(E570C) expression construct.
(B) Cartoon representation of the mouse SMC hinge, with mSmc1 shown in red and mSmc3 in blue. Homologous residues substituted with cysteines for *in vitro* cross-linking of the yeast SMC dimer have been represented as yellow sticks. [PDB: 2WD5. Image generated using The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.]
(C) Western blotting confirmed the presence of both SMCs in the WCE.
(D) His-tag affinity purification from 5 litres of induced culture gave a large, sharp absorbance peak.

[Figure 12 continued on next page.]

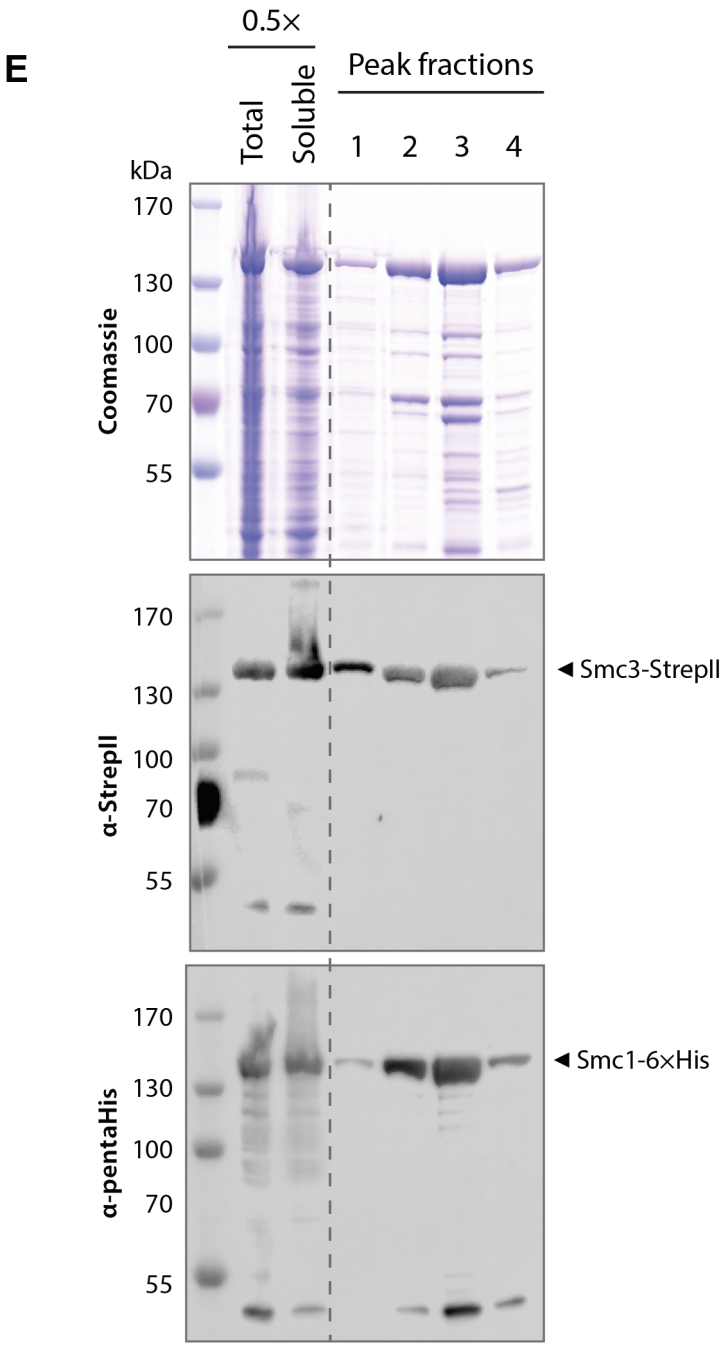

Figure 12. Substitution of E570 in Smc3 and L639 in Smc1 with cysteine residues does not affect the two proteins' co-purification.

(E) Fractions marked in (D) with blue rings contained both SMCs, as determined by SDS-PAGE and Western blotting.

amino acid substitution in both SMCs to create a construct containing Smc3(E570C)-StrepII followed by a RBS and Smc1(L639C)-6×His (Figure 12A-B). Except for the two cysteine changes, the cross-linkable construct was identical to the one used for the SMC dimer purifications. Indeed, induction of expression was also very similar (Figure 12C) and the sharp absorbance peak generated as a result of a larger-scale His-tag purification (5 litres of induced culture) contained both proteins of interest (Figure 12D-E).

The thiol-reactive cross-linkers of choice were dibromobimane (dBBr) and bis-maleimidoethane (BMOE), which can cross-link across distances of up to 5-6 Å and 8 Å, respectively. Purified material was split into separate reactions and each incubated with DMSO (negative control) or one of the two chemical cross-linkers. The reaction was then terminated by boiling the samples in loading buffer and the presence of cross-linked species was detected by Western blotting.

Upon addition of dBBr, a high molecular weight band could be detected in the His-tag purified Smc1/Smc3 sample with cysteine substitutions in the hinge (Figure 13A). The appearance of this band was cross-linker-dependent, since there was no high MW product in the DMSO-only control sample. Western blotting confirmed the presence of both Smc1-6×His and Smc3-StrepII in the high MW band, demonstrating that it was indeed the result of chemical cross-linking specifically between the two proteins in question. A similar pattern was observed for the StrepII-tag purified SMC dimer: incubation with either dBBr or BMOE induced the appearance of a high MW species (Figure 13B). Western

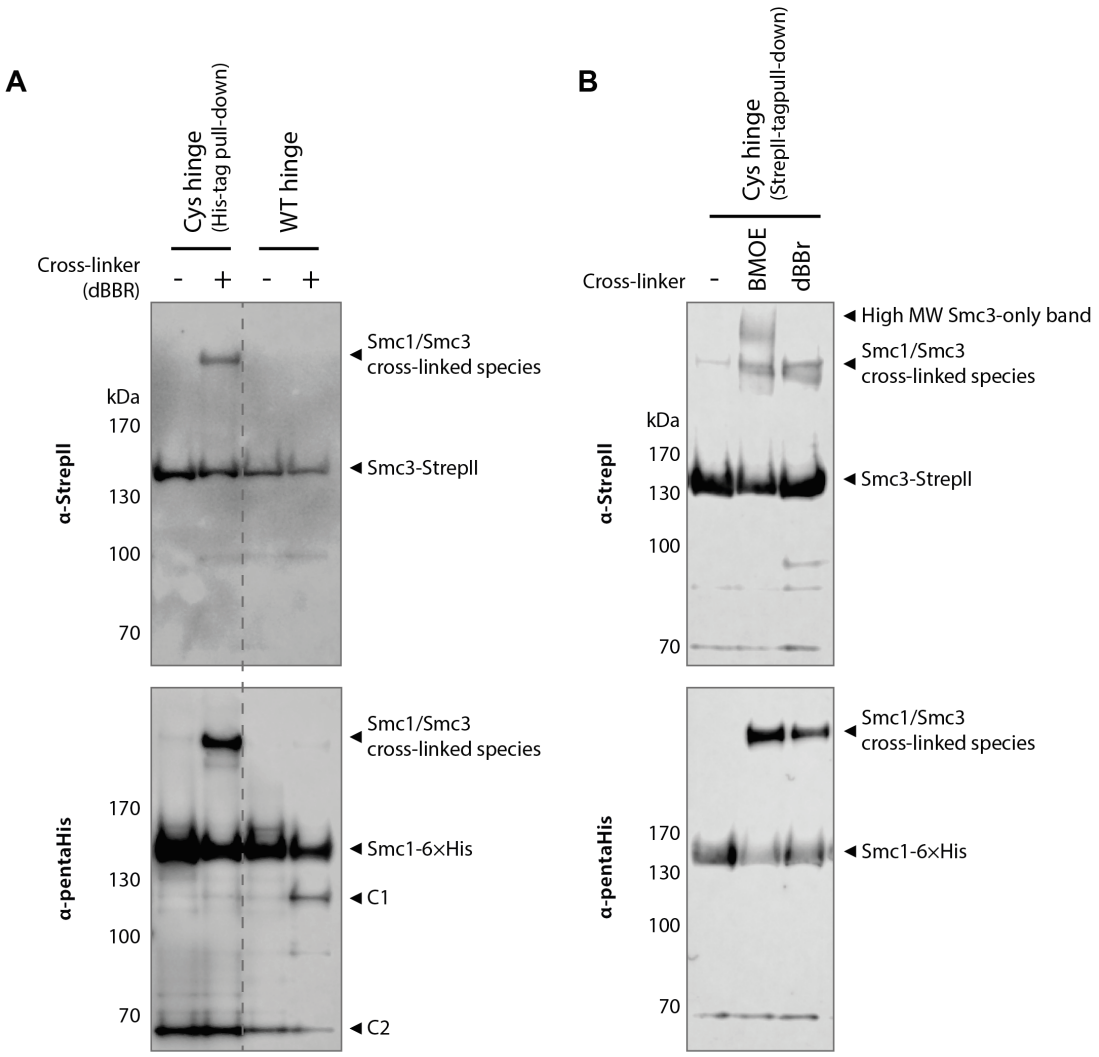

Figure 13. Both SMC purifications result in a functionally assembled Smc1/Smc3 heterodimer, as demonstrated by chemical cross-linking of the hinge interface.

(A) The His-tag affinity purified, full-length SMCs could be chemically cross-linked in a reaction dependent on the substitution of specific residues with cysteines in both SMC hinge domains and the presence of dBBr. ['WT hinge': wild-type Smc1/Smc3; 'Cys hinge': Smc1(L639C)/Smc3(E570C).]
(B) StrepII-tag purified Smc1/Smc3 dimers could be cross-linked at their hinge interaction interface using both dBBr and BMOE. [dBBr: dibromobimane; BMOE: bis-maleimidoethane.]

Note: An extra band marked 'C1' in (B) appeared upon addition of cross-linker in the absence of hinge cysteine residues. This was presumed to be the result of a cross-linked species that contained contaminant marked 'C2' in the sample, as the 'C2' band decreased in intensity upon emergence of band 'C1'.

blotting confirmed the presence of both SMCs in this band, which was a good indication that both purified full-length Smc1/Smc3 dimers were interacting physiologically at the hinge interface.

The appearance of the cross-linked species was also dependent on the presence of the pair of cysteine residue substitutions in the hinge, as no band of such high molecular mass was detected in the His-tag purified WT hinge. That said, to rigorously demonstrate the requirement for both cysteine residues, one would need to include an additional two constructs in the cross-linking experiment: WT Smc1 with Smc3(E570C) and Smc1(L639C) with WT Smc3. This way, the possibility that a cysteine in one SMC cross-links non-specifically to the other SMC would be excluded.

While the dBBr experiment using the StrepII-tag purified sample looked remarkably similar to that observed for the His-tag dimer, the StrepII-tag purified BMOE sample contained an extra, high MW band not present in the dBBR lane (marked in Figure 13B). This species reacted only against the α-StrepII antibody, implying it contained Smc3-StrepII but not Smc1-6×His. An explanation for the presence of this extra band relates to the fact that WT Smc3 contains three cysteine residues: one in the hinge region and two obscured in the head domain, apparently unable to react with the cross-linker (Gligoris et al., 2014). If there were free Smc3 molecules in the eluate and a proportion of them homodimerised via interactions in their hinge domains, the resulting product would be of a similar shape to the Smc1/Smc3 cross-linked species and, hence, would co-migrate with it. However, the observed Smc3-only high MW band ran

slower through the gel, so it likely contained a differently shaped cross-linked species. For example, were one Smc3 hinge to cross-link with another Smc3 head, the resulting molecule would likely migrate separately from the Smc1/Smc3 species. Such an Smc3/Smc3 species could only have been formed if some of the original protein had been denatured. Since BMOE can react with thiol groups up to 2 Å further away than dBBr, it is not inconceivable that a hinge cysteine residue and a head domain cysteine from a denatured Smc3 could have come in close enough proximity for a more 'promiscuous' cross-linker like BMOE to straddle. The above logic is based on the existence of free Smc3 in the sample: the most plausible explanation is that Smc3 is expressed at higher levels than Smc1 and this possibility was investigated in the following section.

As briefly mentioned, the appearance of the high MW band containing both Smc1 and Smc3 in a process dependent on the presence of both cross-linker and specific cysteine pairs implied that material obtained from both purifications of the SMC dimer was functionally assembled at the hinge interface and, hence, likely to behave physiologically. However, it must be noted that these findings give only indirect evidence as to whether the coiled-coils and head domains of the SMCs are also folded physiologically. For cross-linking to occur at the hinge interface, each SMC's hinge domain must be correctly folded. The amino acids that make up the hinge domain are in the middle of the primary sequence of each SMC and are linked to the head domains at the termini by stretches of coiled-coil. Without further rigorous testing, one can therefore only assume that the 'kink' in the sequence induced as a result of the hinge domains assembling correctly is

sustained solely if the coiled-coils (and presumably also the heads) are also constructed correctly.

Calculated hinge cross-linking efficiencies demonstrate that the StrepII-tag Smc1/Smc3 eluate is not purely heterodimeric

A cross-linking event reliant on a bi-functional chemical reacting with two thiol groups, each on a different molecule, in almost all cases is not 100% efficient. Amongst others, the following can all affect the amount of cross-linked product observed: obscured access to or an excessive distance between the cross-linkable cysteine residues due to protein flexibility in solution, only one of the cross-linker/thiol reactions going to completion, and the cross-linker being depleted by reacting preferentially with non-cross-linkable cysteine residues. Typically, during an efficient cross-linking reaction, depletion of the cross-linkable substrate will be observed i.e. most of that protein's signal 'migrates' into the high MW cross-linked band. The percentage of total signal composed of the cross-linked species is the defined as the cross-linking efficiency for a given reaction.

Cross-linking efficiency of the SMC hinge interface can also be used to assess the complex's stoichiometry. It is an implicit fact that for every Smc1 molecule that shifts into the high MW band following hinge interface cross-linking, one Smc3 molecule must be depleted. Therefore, in a stoichiometric complex, the number of molecules cross-linked must represent the same percentage of the total

number of molecules for both SMCs. In other words, if the purified Smc1/Smc3 sample is purely heterodimeric with a 1:1 ratio, the cross-linking efficiency value should be identical for both proteins. Should the purified material have a surplus of one SMC, its cross-linking efficiency will be lower than that of its less abundant binding partner, as there will be more non-cross-linked material.

Fiji image analysis software was used to calculate the cross-linking efficiency for both SMC dimer purifications (His-tag and StrepII-tag). This was done by analysing the intensity values for the semi-quantitative chemiluminescent Western blots presented in Figure 13 that had been obtained using an ODYSSEY Fc Imaging System (LI-COR Biosciences). More precisely, the cross-linked species' intensity value was divided by the summed intensities of the cross-linked and non-cross-linked bands within a single sample lane (Schindelin et al., 2012). The cross-linking efficiency was estimated to be approximately 42 % for both SMCs when using material obtained from the His-tag affinity purification (Table 2). This was not the case for the StrepII-purified dimer, where the cross-linking efficiencies for Smc1 and Smc3 were 53.4% and 19.5%, respectively.

Table 2: Cross-linking efficiency of each SMC hinge subunit varied between His-tag and StrepII-tag purification methods.

Purification/ cross-linker	Cross-linking efficiency	
	Smc1	Smc3
His-tag/dBBr	41.3 %	42.0 %
StrepII-tag/dBBr	53.4 %	19.5 %
StrepII-tag/BMOE	77.8 %	24.7 %

The cross-linking efficiency values imply that the SMC dimer obtained from His-tag purification is heterodimeric with a stoichiometric ratio. Material obtained using the StrepII-tag protocol is partially heterodimeric, but also contains an excess of Smc3-StrepII. In fact, the efficiency data for both cross-linkers suggest that Smc3 is present in an approximately 3-fold excess to Smc1.

Taken together, the data obtained throughout the purification and characterisation of the Smc1/Smc3 dimer have three important conclusions:

(i) *S. cerevisiae* Smc1 and Smc3 can be co-purified as full-length proteins using a bacterial expression system. The complex obtained is heterodimeric with a stoichiometric ratio and the hinge interface is functionally assembled. This is a tool of great value to the field, as recombinant yeast SMCs can now be produced at high yields relatively quickly and easily.

(ii) The SMC dimer is soluble and properly assembled in a 'mild' buffer and at a physiological salt concentration (150 mM). It is also not a marked target for non-specific proteolysis. These features are critical for the ultimate goal of using the purified yeast cohesin complex in *in vitro* assays: the reconstituted complex must be stable and functional in a physiological buffer.

(iii) The purification system described in this thesis results in an excess of Smc3 being expressed, compared to Smc1. Using His-tag affinity purification, which pulls on the lesser expressed component Smc1, results in a relatively impure elution but a heterodimeric stoichiometric Smc1/Smc3 complex. If needed, the

purity of the elution could doubtless be improved by size exclusion or ion exchange chromatography.

Seeing that the yeast SMC heterodimer can be reliably purified, the next step was to attempt to complete the cohesin trimer with Scc1.

Soluble, recombinant Scc1 can be recovered by His-tag purification and the purity of the eluate is boosted by anion exchange chromatography

As mentioned earlier, one of the main concerns previously encountered in the lab regarding yeast Scc1 purification was that the recombinant version was mostly insoluble, largely a consequence of the unstructured stretches in its sequence. This study was the first to use a version of yeast *SCC1* codon-optimised for *E. coli*, which was expected to boost protein expression but not affect its solubility. However, during the creation of the Scc1/Smc1 construct discussed in Figure 9, it was noted that Scc1-6×His alone in the pET28a vector (Figure 14A) was expressed at levels high enough to detect in the WCE (Figure 14B). Western blotting confirmed the presence of a band of the expected size for full-length Scc1. It was therefore decided to pursue this further and test whether any of the synthesised material was soluble.

In lieu of a formal solubility test, 1.6 litres of induced bacterial culture was used to carry out the first few steps of a His-tag purification. Samples used to assess protein solubility were taken following lysis in a French press prior to and after

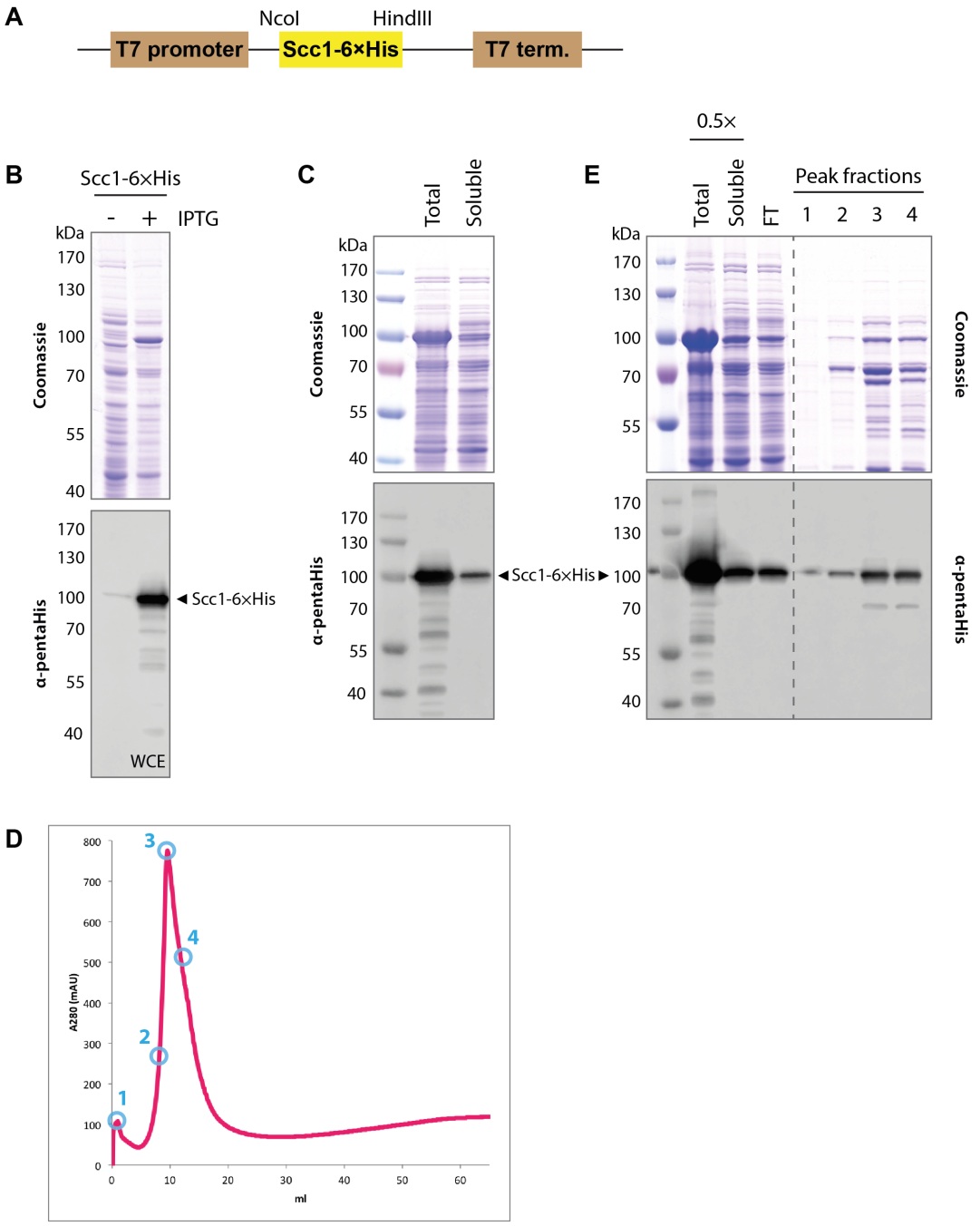

Figure 14. Recombinant yeast Scc1-6×His is partially soluble and can be affinity purified from *E. coli*. Sample purity can be greatly enhanced by anion exchange chromatography.

(A) Representation of the Scc1 purification construct.
(B) Induction of the expression strain resulted in expression of Scc1 to levels high enough to detect in the WCE. Western blotting confirmed the identity of the protein.
(C) A small-scale solubility test revealed that most of the expressed recombinant Scc1 was insoluble.
(D) His-tag purification of the Scc1 construct gives a single, sharp absorbance peak.
(E) The His-tag eluate contained full-length Scc1, as determined by Western blotting, but Coomassie staining revealed that it also contained a large proportion of protein contaminants.

[Figure 14 continued on next page.]

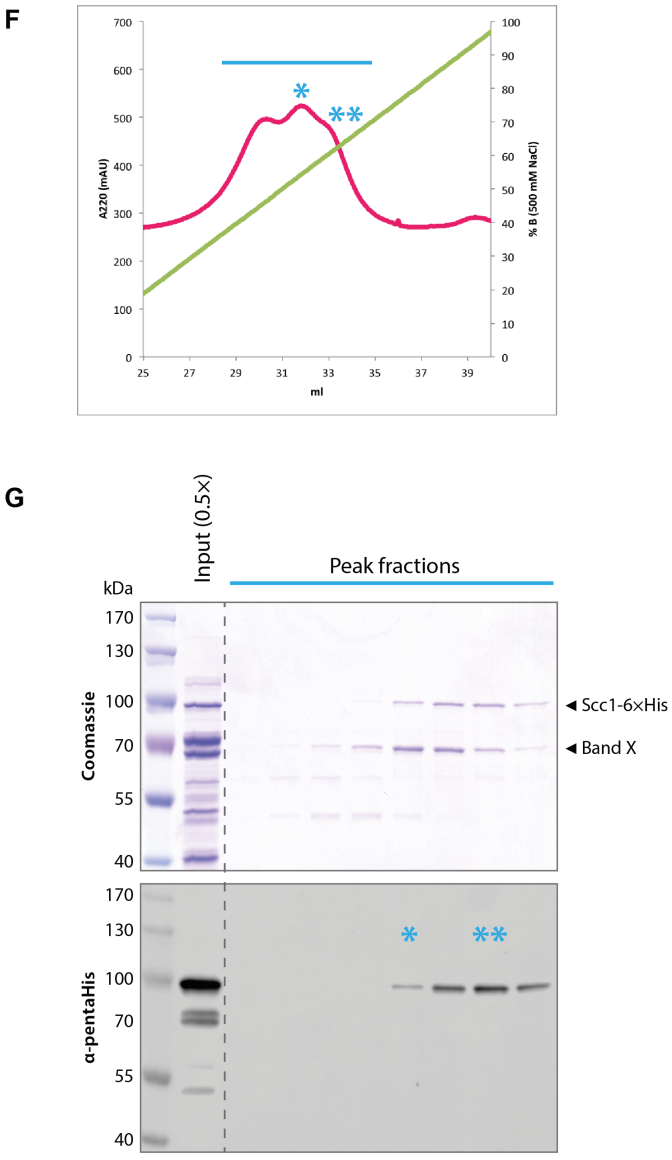

Figure 14. Recombinant yeast Scc1-6×His is partially soluble and can be affinity purified from *E. coli*. Sample purity can be greatly enhanced by anion exchange chromatography.

(F) Passing the His-tag elution through an anion exchange column resulted in a broad, multimeric absorbance peak at 220 nm.
(G) The latter half of the broad absorbance peak contained two protein bands, one of which was confirmed to be full-length Scc1-6×His by Western blotting. Asterisks indicate which peaks correspond to which lanes on the gel. [Mass spectrometry was later used to show that 'Band X' is a bacterial contaminant.]

ultracentrifugation. Quantification of the resultant Western blot revealed that only about 30% of the total material was soluble in the lysis buffer used (Figure 14C). This was nonetheless deemed to be a promising amount of soluble Scc1 and the strategy of combining it with the successfully purified SMC dimer worth pursuing.

A His-tag purification was performed using the soluble material obtained from 3 litres of bacterial culture expressing the Scc1-6×His construct. The elution produced a relatively large, sharp absorbance peak (Figure 14D) that contained the protein of interest and a faint band containing a small amount of N-terminal cleavage product (Figure 14E). Similarly to the His-purified SMC dimer, there was very little depletion of His-tagged material. Also as noted for previous His-tag purifications, Coomassie-staining revealed that the Scc1 elution contained a large proportion of protein impurities. The most prominent contaminants were two bands around 70 kDa, the faster migrating one of which had been identified by mass spectrometry as glutamine-fructose-6-phosphate aminotransferase in the SMC dimer sample.

In order to reconstitute a biochemically sound cohesin complex, there should be few, if any, contaminants present in the purified material. Size exclusion chromatography is a good choice for removal of small-molecule impurities (for example, imidazole that had been used to elute the protein) and contaminating proteins/peptides, if their sizes are not too similar to the protein of interest. The principle of size exclusion chromatography relies on a molecule's size and shape (e.g. globular vs. extended), whereby smaller, globular molecules are retarded by

the matrix more than larger or extended ones. Thus, the 'spindly' SMC dimer (~280 kDa) should run through a size exclusion column slightly faster than a globular protein of 280 kDa. The complex would be also separated from the much slower-running ~70 kDa contaminants detected in the His-tag elution, unless these were part of a larger complex prior to denaturation.

While recombinant Scc1 has a molecular mass of 64 kDa, it was found to run alongside the 100 kDa marker on a polyacrylamide gel, presumably because of a charge effect (Rath et al., 2009). However, since size exclusion chromatography is not dependent on charge, the ~70 kDa bacterial contaminants detected after His-tag purification would most likely elute off the column together with Scc1. Therefore, size exclusion chromatography could not be used to further purify the Scc1 His-tag elution. Instead, ion exchange chromatography was used, which is a technique that separates proteins based on their net surface charge. Since proteins are amphoteric, their surface charge changes depending on the surrounding pH in a manner unique to each protein: this feature is used as the principle of protein separation (Eith et al., 2001).

The pH at which a protein's net charge is zero is called its isoelectric point (pI). At pH values above their pI, proteins are capable of binding positively charged surfaces or 'anion exchangers.' The His-tag elution buffer has a pH of 7.5, so the Scc1-6×His in the eluate should be capable of binding an anion exchange column since the protein's pI is 4.64, as calculated using the ExPASy ProtParam tool (Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., 2005). Proteins capable of interacting with the positively-charged matrix were

competitively eluted using an increasing concentration of chloride ions to generate a cluster of small, poorly resolved peaks of absorbance at 220 nm (Figure 14F). The 220 nm wavelength was used to illustrate the pattern observed due to its increased sensitivity.

Fractions throughout the wide ion exchange elution peak were analysed by SDS-PAGE to reveal that the latter half contained two protein bands of 100 kDa and 70 kDa. The larger of the bands was determined to be full-length Scc1-6×His by Western blotting (Figure 14G). To unequivocally identify the two bands, both the Scc1 band and the 70 kDa 'Band X' were analysed by mass spectrometry, in collaboration with Dr Benjamin Thomas (Central Proteomics Facility, Sir William Dunn School of Pathology, University of Oxford). As expected, the band running at 100 kDa was indeed composed of mainly yeast Scc1, while the smaller one was the bacterial enzyme glutamine-fructose-6-phosphate aminotransferase. Despite this single bacterial contaminant, the purity of the His-tag elution was greatly boosted by anion exchange chromatography, as is apparent when comparing the 'input' to the elution lanes in Figure 14G. This Scc1 sample, along with His-tag purified SMC heterodimers were therefore used for the following trimer reconstitution experiment.

Scc1 co-elutes with the SMC heterodimer from a size exclusion column

As shown before, the His-tag purified SMC heterodimer was functionally assembled at the hinge interface, but little could be said regarding its coiled-coils

and head domains. One way to show physiological folding of the SMCs would be to demonstrate that the dimer is capable of binding Scc1, since this interaction depends on the presence of both Smc1 and Smc3 head domains (and part of the coiled-coil in the case of the Scc1/Smc3 interface) (Gligoris et al., 2014; Haering et al., 2004).

One convenient method to determine whether proteins form a complex is by testing whether the migration of individual components though a size exclusion column is retarded by addition of potential binding partners (Rowland et al., 2009). If this is the case, a larger proteinaceous species must be being formed: its size can be estimated by comparison to gel filtration standards (Figure 15A) and its composition determined by standard proteomic techniques. Following this logic, addition of purified Scc1 to the SMC heterodimer should decrease the retention time of the dimer in a gel filtration column if the three proteins interact. In other words, formation of the cohesin trimer would shift the absorbance pattern generated by the SMC dimer alone to the left (smaller elution volume). It is important to note that it is difficult to distinguish a physiological interaction from a non-specific one using this technique: proteins interacting with each other specifically would run through a gel filtration column comparably to those binding non-specifically but making up a similarly sized and shaped mass.

The protein concentrations of the wild-type, His-tag purified SMC dimer and the anion exchange-purified Scc1-6×His were determined. Each purified sample was then split into two: one half was passed through the size exclusion column on its

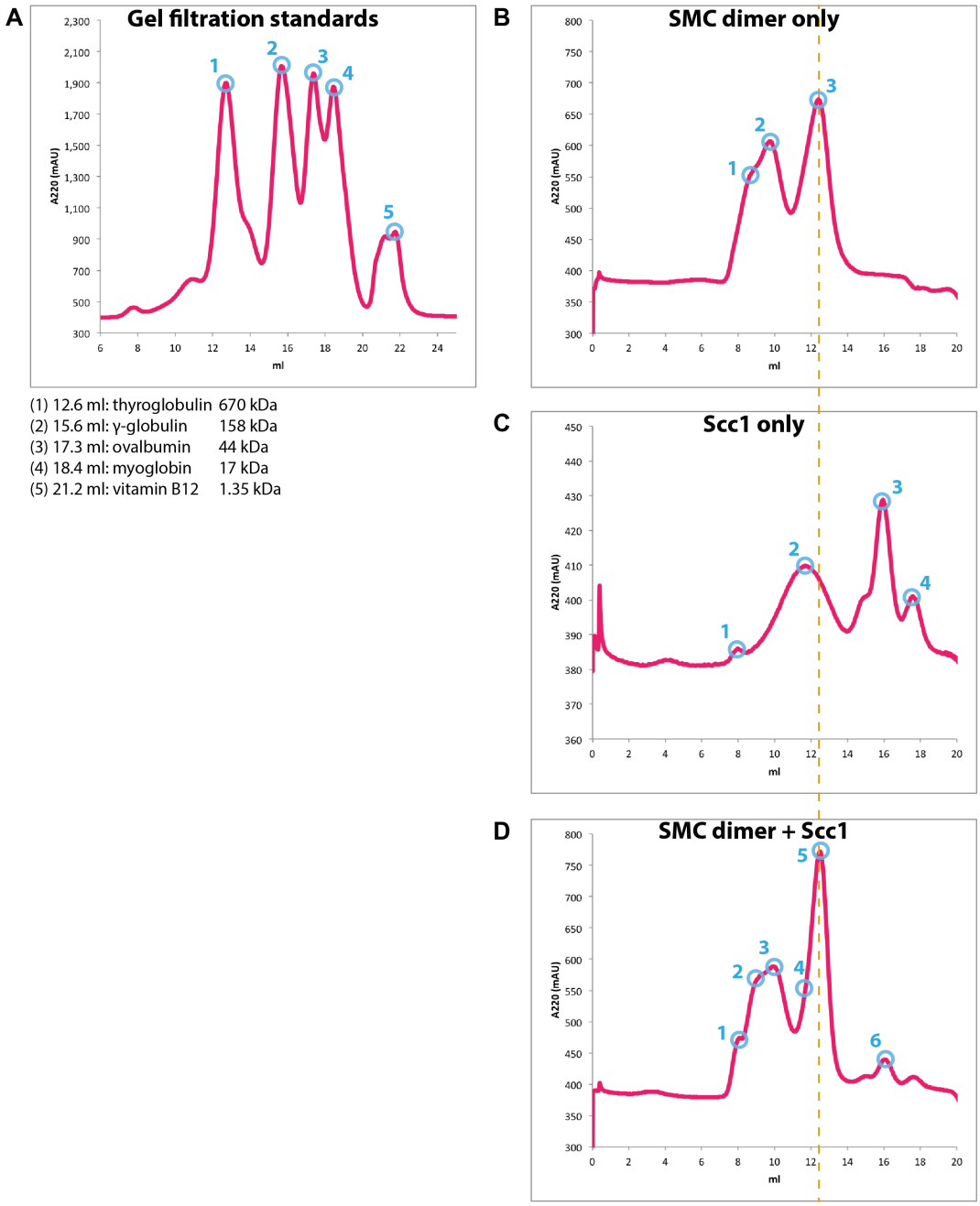

Figure 15. Scc1 co-elutes from a gel filtration column with the SMC heterodimer.

(A) Gel filtration standards (Bio-Rad), as resolved under the same conditions used for all cohesin samples.
(B)-(D) Chromatograms of absorbance at 220 nm of the three samples analysed by size exclusion chromatography: (B) 'SMC dimer only', (C) 'Scc1 only', and (D) 'SMC dimer + Scc1.' Blue rings indicate which fractions were analysed by SDS-PAGE. The dashed orange line marks the same elution volume in all samples for comparison of the peaks' position.

[Figure 15 continued on next page.]

E

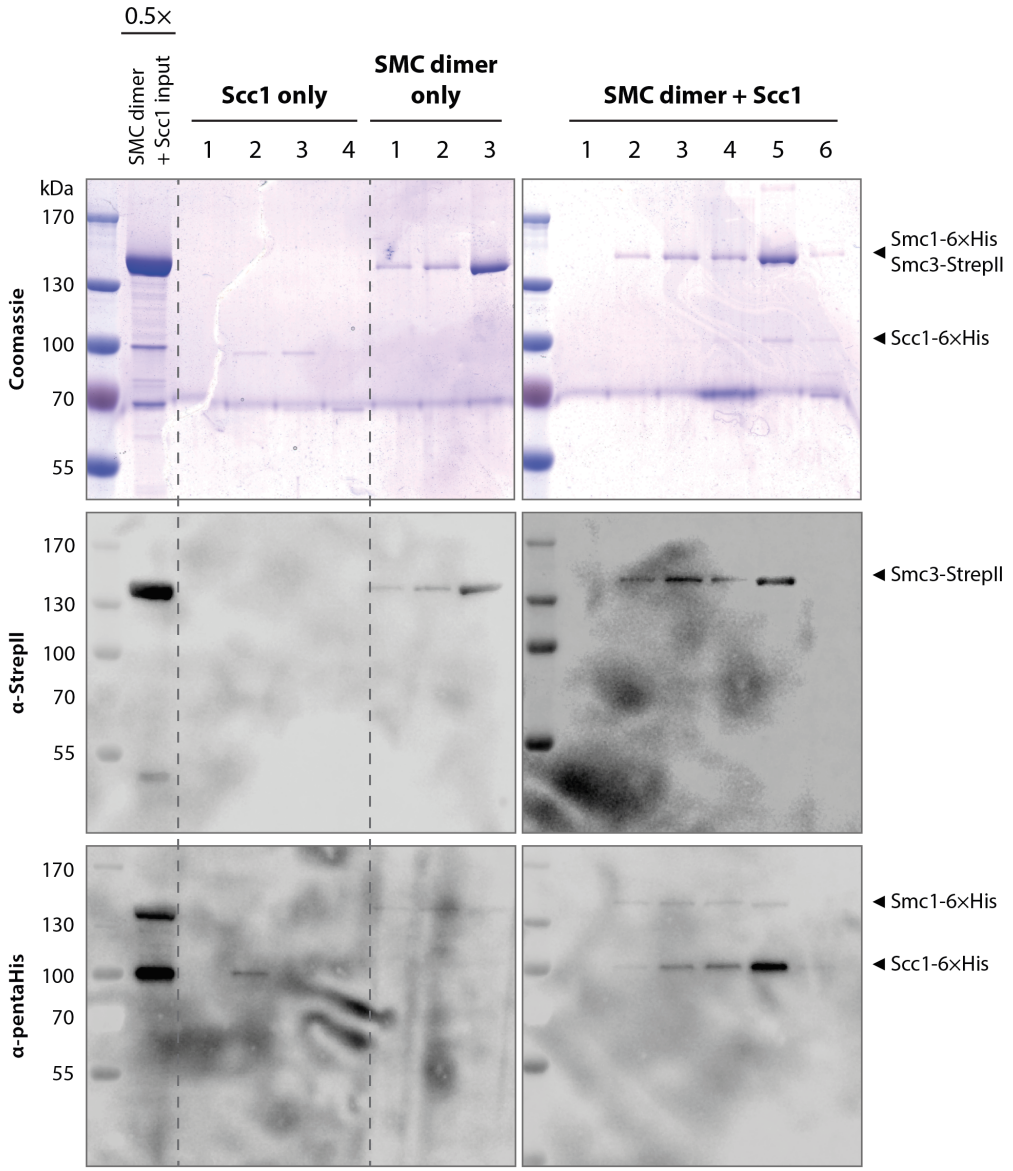

Figure 15. Scc1 co-elutes from a gel filtration column with the SMC heterodimer.

(E) Western blotting confirmed the presence of both SMCs in all absorbance peaks of the chromatogram of the 'SMC dimer only' sample, with the bulk of material present in the largest peak at ~12.5 ml. When gel filtrated alone, Scc1 was present predominantly in the wide peak at ~11.5 ml. The 'SMC dimer + Scc1' sample contained both SMCs in the two major peaks observed. Scc1 in this sample could be detected in the small absorbance peak at ~9 ml, but most of it was present in the ~12.5 ml peak.

own while the other half was used to attempt cohesin trimer reconstitution (the 'SMC dimer + Scc1' or 'trimer' sample). This was done by mixing the SMC dimer and Scc1 such that there was a 2:1 excess of Scc1 to ensure that it was not limiting the reaction. The Superose 6 10/300 GL gel filtration column used in the following experiments was capable of separating complexes between 5 and 5,000 kDa. The recombinant cohesin trimer was predicted to have a molecular mass of ~350 kDa but its 'extended' architecture meant that it would most likely elute in a smaller volume than would be expected for a globular complex of the same size.

Size exclusion chromatography of the SMC dimer revealed that the sample eluted from the gel filtration column in two distinct peaks (Figure 15B). The first peak was broad, spanning elution volumes from ~8 to ~10 ml and likely composed of two unresolved peaks; the second peak at ~12.5 ml was larger and sharper. Western blotting analysis of samples from all three peaks revealed that both SMCs were present throughout, with most protein detected in the peak at ~12.5 ml (Figure 15E). Coomassie staining demonstrated that the eluted SMC dimer was free of any non-specific protein contaminants, unlike the original His-tagged elution (see Figure 11D).

It is important to note that while size exclusion chromatography can relatively accurately estimate the MW of globular proteins, the migration of extended molecules like the SMC dimer is not easily predictable. It is hence not unexpected that the ~280 kDa SMC dimer eluted at a volume corresponding to or below that observed for a 670 kDa globular protein (12.6 ml; Figure 15A). That said, both

peaks observed in the SMC sample contained the proteins of interest, indicating that the originally purified material was heterogeneous. One explanation would be that the larger MW species was made up of SMC multimers while the smaller MW peak contained only dimers (or at least smaller oligomers): this possibility should be addressed by MALS.

Gel filtration of the 'Scc1 only' sample showed that it migrated through the column in at least three distinct populations: a broad peak at ~11.5 ml and two sharp peaks at ~16 and ~17.5 ml (Figure 15C). Only the first peak was shown by Western blotting to contain Scc1-6×His (Figure 15E). Coomassie-staining of samples from the remaining peaks did not show a distinct signal, so they were assumed to contain nucleic acid contaminants. As with the SMC dimer, the Scc1 sample did not run through the matrix like a globular protein of the same size would (between 16 and 18 ml; Figure 15A). This could be partially explained by Scc1's unstructured, flexible conformation, but the possibility that the purified material oligomerises to form a larger complex cannot be excluded.

Addition of Scc1 to the SMC dimer resulted in an absorbance peak pattern similar to that of the 'SMC dimer only' sample (Figure 15B). Two extra absorbance peaks were detected: one at an elution volume of ~8 ml and one at around 16 ml, but these did not contain detectable protein material. It was found that samples throughout the two major peaks observed across ~8-10 ml and at ~12.5 ml contained all components of the cohesin 'trimer': Smc1-6×His, Smc3-StrepII, and Scc1-6×His. Smc3 and Scc1 were most abundant in the later-eluting material; Smc1 was present in small amounts in both peaks. Three fractions

spanning the ~12.5 ml absorbance peak were selected for the next step of the investigation, as they contained the largest amount of correctly-staining protein components.

The observation that Scc1 co-eluted from the gel filtration column together with the SMC dimer is a good indication for complex assembly. That being said, addition of a 64 kDa Scc1 molecule to the SMC dimer should have shifted the absorbance peak observed at ~12.5 ml in the 'SMC dimer only' sample towards the left i.e. a smaller elution volume. This was not the case: the dashed orange line in Figure 15 crosses the largest peaks at the same position in both panels B ('SMC dimer only') and D ('SMC dimer + Scc1'). It should also be noted that the vertical orange line crosses the major Scc1 peak in the 'Scc1 only' sample reasonably close to the peak's apex. It is therefore conceivable that a substantial proportion of Scc1 eluted at a volume of ~12.5 ml regardless of the presence of the SMC dimer. In order to exclude this possibility and demonstrate functional cohesin trimer assembly, the cross-linking experiments described below were performed.

Functional assembly of the 'trimer' fraction cannot be confirmed by chemical cross-linking between Smc3(S1043C) and Scc1(C56)

In order to test whether the interaction detected in the 'SMC dimer + Scc1' sample was indeed physiological (i.e. consistent with what has been observed *in vivo*), a similar approach as was used to demonstrate a functional SMC dimer

hinge interaction was taken. Namely, site-specific crosslinking of either of the SMC/Scc1 interfaces should determine whether this part of the 'timer' is assembled correctly, as the cross-linking reaction should only occur within a specific, short range of distances. Detection of a high MW species containing both the SMC in question and Scc1 following incubation with chemical cross-linker would be a good indication that a physiological interaction is taking place.

The residue pair of choice made use of an endogenous cysteine in Scc1 (C56) that has been shown to be capable of cross-linking to the Smc3(S1034C) substitution *in vivo* (Gligoris et al., 2014). While other cross-linkable cysteine pairs are available both in the Smc3/Scc1 and Smc1/Scc1 interfaces, this approach was deemed most adequate to the timeframe available. In fact, it only required a single round of SDM to create and express one new construct, the Smc3(S1043)-StrepII/Smc1-6×His dimer. As before, a 'bi-cistronic' construct was designed for introduction into the pET28a vector (Figure 16A).

Similarly to the WT SMC dimer, the Smc3(S1043)-StrepII/Smc1-6×His construct ('Cys SMC dimer') was expressed at levels just high enough to detect in the WCE by Coomassie-staining (Figure 16B). The induction of both target proteins was also confirmed by Western blotting. His-tag affinity purification of material from 5 litres of culture resulted in a broad absorbance signal from two unresolved peaks (Figure 16C) which both contained Smc1 and Smc3 (Figure 16D). Coomassie-staining of the whole protein content revealed that the elution was much cleaner than observed previously for SMC dimer His-tag purifications. This was most likely attributed to a higher concentration of imidazole in the wash

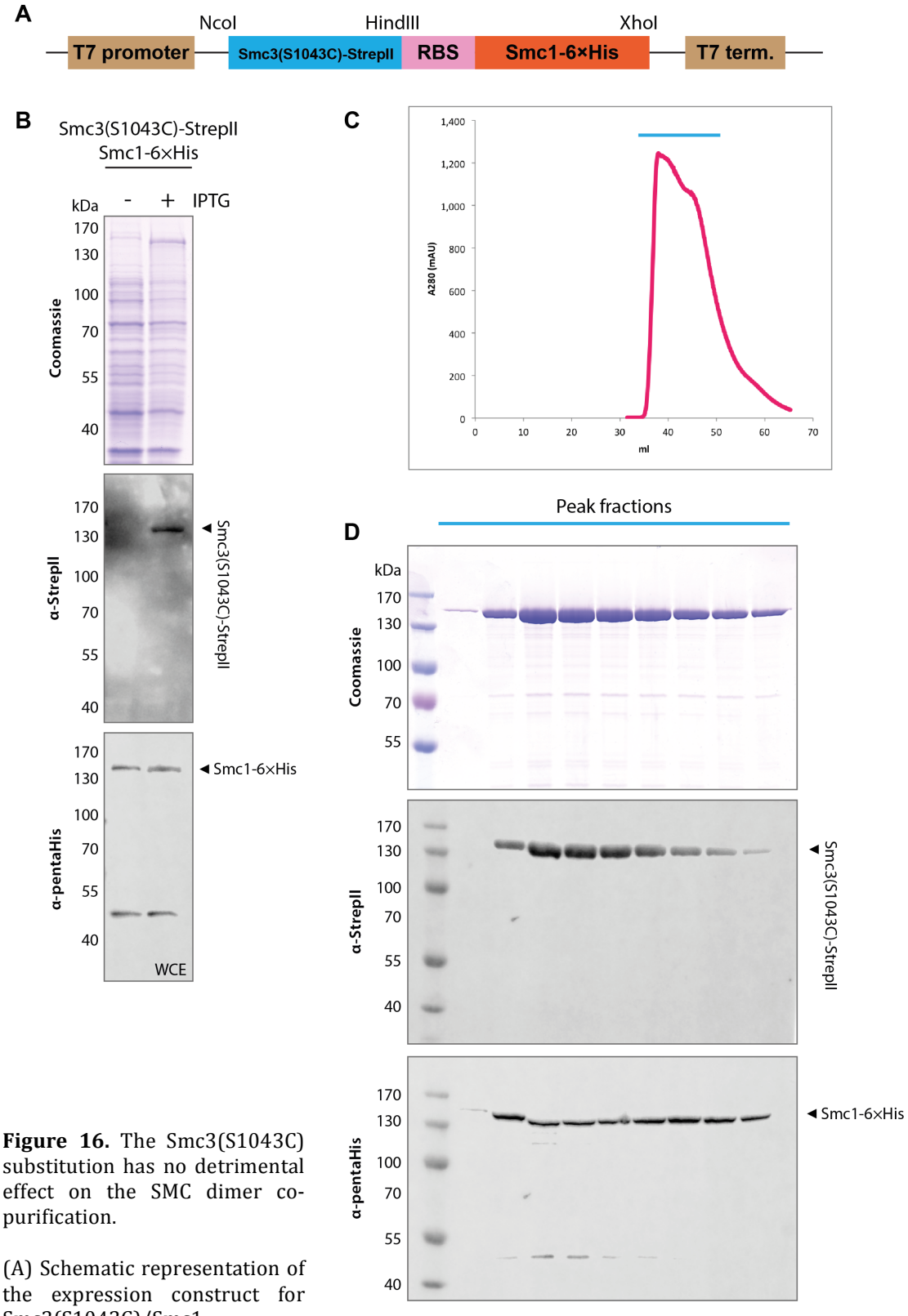

Figure 16. The Smc3(S1043C) substitution has no detrimental effect on the SMC dimer co-purification.

(A) Schematic representation of the expression construct for Smc3(S1043C)/Smc1.

(B) Incubation with IPTG induced the expression of both SMCs, as determined by Western blotting. The proteins of interest were expressed to levels high enough to detect in the WCE.

(C) His-tag affinity purification from 5 litres of bacterial culture generated two large, unresolved peaks of absorbance at 280 nm. [Note: Peak observed later than usual due to omitted pump wash.]

(D) All fractions throughout the breadth of the peaks contained both Smc1 and Smc3, as detected by Western blotting and Coomassie staining. The elution was much cleaner than previously noted for His-tag purifications, likely due to the higher concentration of imidazole used in the wash buffer (25 mM).

buffer (25 mM) that competitively displaced non-specifically bound material.

The size exclusion chromatography experiment using the 'Cys SMC dimer' was performed in the same manner as for the WT sample. The gel filtration traces obtained were remarkably similar to the WT samples, demonstrating that introduction of a cysteine residue at Smc3's position S1043 was not detrimental to the dimer's purification properties or to the observed co-elution of Scc1 with the SMC dimer (Figure 17). Fractions covering the major peak at ~12.5 ml of the 'Cys SMC dimer + Scc1' chromatogram were chosen for the cross-linking experiment, along with samples from the 'WT SMC dimer + Scc1' gel filtration. His-tag purified (not gel filtrated) SMC dimers with either WT hinges or cross-linkable hinges were run alongside as negative and positive controls, respectively.

Gel filtrated samples of Scc1 with either an SMC dimer containing WT Smc3 or Smc3(S1043C) were incubated with dBBr prior to protein detection by Western blotting. All input proteins could be detected on the immunoblot, but the amount of total protein in the gel filtrated samples was much lower than for the hinge controls (Figure 18). While a cross-linked band could be detected in the SMC hinge positive control sample (cysteine pair and dBBr present), no such high MW species could be detected in any of the 'trimer' lanes. There are several explanations for this, as discussed below:

(i) Perhaps the low protein levels in the input samples were simply obscuring the result and cross-linking within the 'trimer' was actually taking place. While

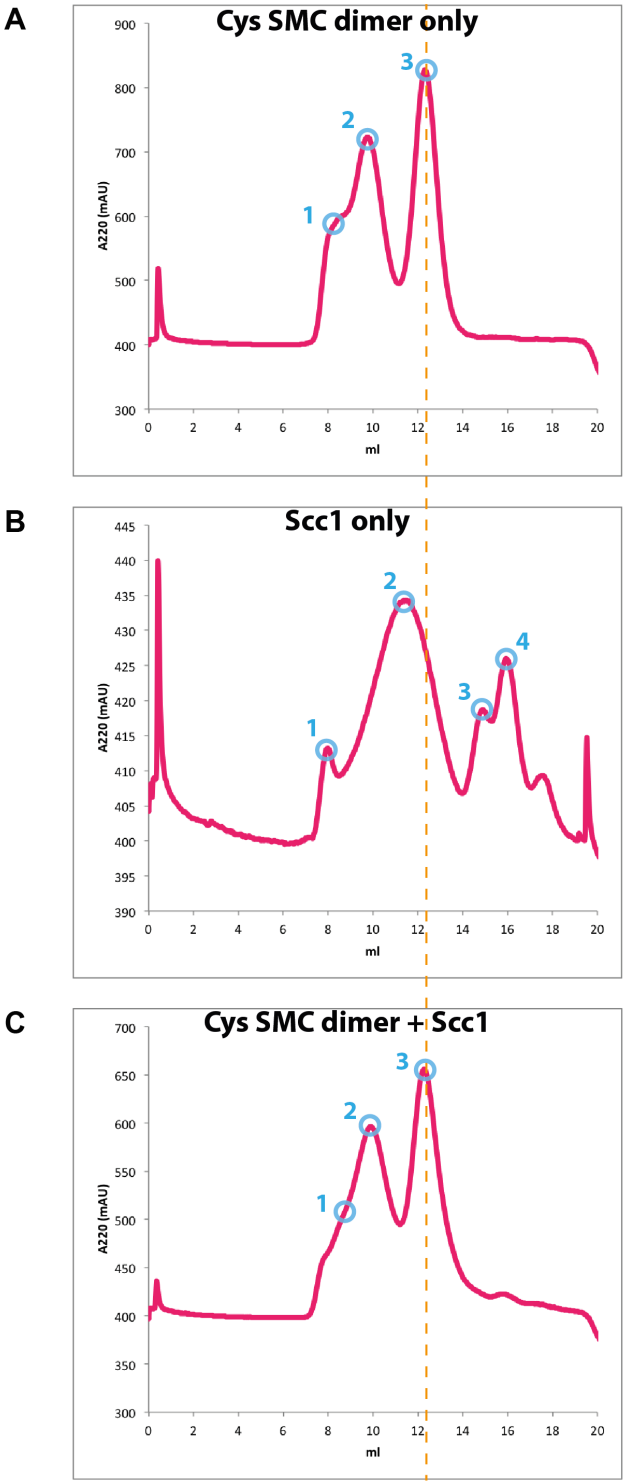

Figure 17. Introduction of a cysteine residue at position S1043 in Smc3 does not affect the ability of Scc1 to co-elute with the SMC dimer from a gel filtration column.

(A)-(C) Chromatograms of absorbance at 220 nm of the three samples analysed by gel filtration: (A) 'Cys SMC dimer only', (B) 'Scc1 alone', and (C) 'Cys SMC dimer + Scc1.' Fractions analysed by SDS-PAGE are marked by blue rings. The dashed orange line crosses the same elution volume in all chromatograms to enable comparison of the peaks' position.

[Figure 17 continued on next page.]

D

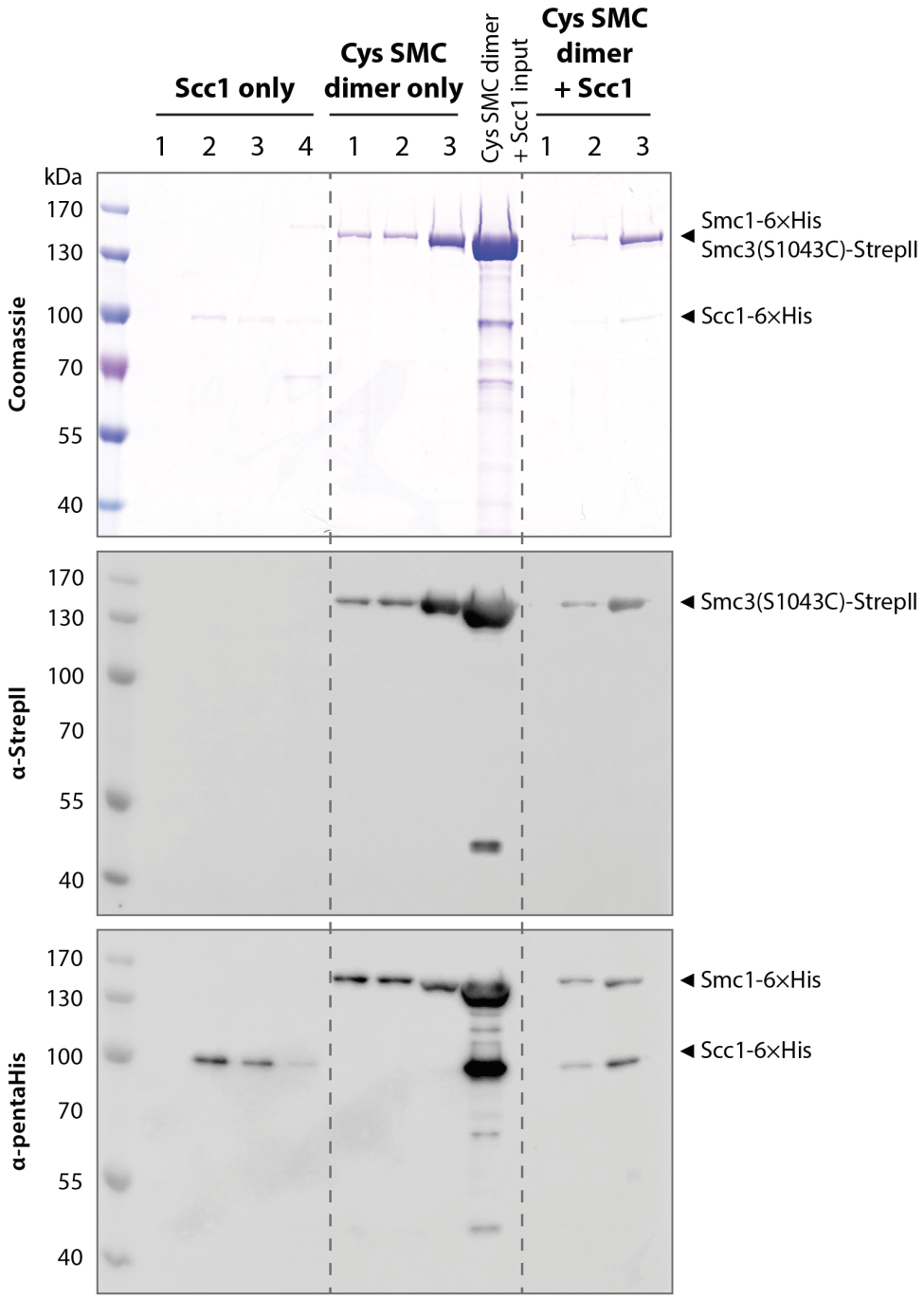

Figure 17. Introduction of a cysteine residue at position S1043 in Smc3 does not affect the ability of Scc1 to co-elute with the SMC dimer from a gel filtration column.

(D) The 'SMC dimer only' absorption spectrum contained both SMCs in all fractions analysed by Western blotting, with most material being detected in the largest peak at ~12.5 ml. Gel filtration of 'Scc1 only' determined that Scc1 was present predominantly in the peak at ~11.5 ml. The 'Cys SMC dimer + Scc1' sample contained all three proteins in peaks at ~10 ml and ~12.5 ml, although most material was observed in the latter peak.

the Smc3-StrepII signal was strong when compared to the hinge controls, the remaining two components of the cohesin trimer seemed to be present only in trace amounts (Figure 18). To exclude this possibility, cross-linking should be repeated following immunoprecipitation of the 'trimer': this would both concentrate the proteins and demonstrate whether their interaction is robust enough to resist a co-IP experiment. It is also unclear how efficient the cross-linking at the Smc3/Scc1 interface was under the conditions used. Using a more directly comparable set of controls would address this: ideally, co-expressed Smc3(S1043C)/Scc1 following His-tag purification and gel filtration.

(ii) The simplest explanation for the negative result observed is that the 'trimer' was not interacting functionally and the co-elution of Scc1 with the SMC dimer from a gel filtration column was non-specific. The three proteins may not be binding in a manner that enables Scc1's endogenous C56 to be positioned close enough to Smc3(S1043) for chemical cross-linking, a distance of approximately 5-6 Å for dBBr. Two pieces of evidence give weight to this explanation. Firstly, there was no trace of depletion of non-cross-linked material for the 'trimer' in Figure 18. Even if a cross-linked species is of too low abundance to detect on a Western blot due to inefficient transfer of high MW bands, the depletion of material to make up this band should be discernable, especially given that the Smc3(S1043C)/Scc1(C56) pair has been shown to cross-link efficiently (Gligoris et al., 2014). Secondly, the Smc1-6×His and Scc1-6×His immunoblot signals were not of equal intensities in the co-eluted sample after size exclusion chromatography (Figure 15E). Since both proteins were tagged with the same epitope, the intensity of their signals should be equal if they were present at a

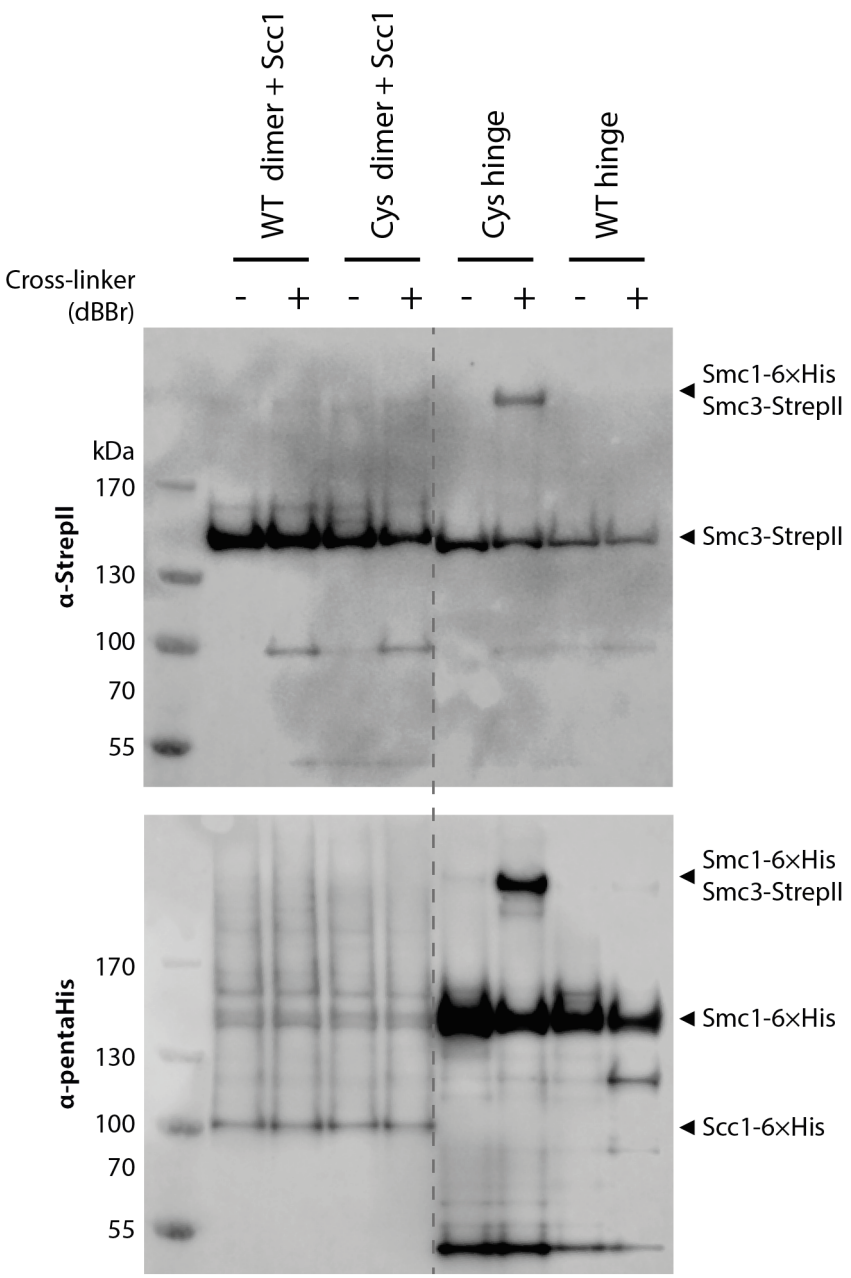

Figure 18. Functional 'trimer' assembly cannot be confirmed by chemical cross-linking between Smc3(S1043C) and Scc1(C56).

Gel filtrated samples of SMC dimer containing either WT Smc3 or Smc3(S1043C) as well as Scc1 were incubated with dBBr prior to protein detection by Western blotting. Control samples for the cross-linking reaction were run alongside: the SMC dimer with WT hinges (negative control) or with cross-linkable cysteine residues in the hinge (positive control).
['WT dimer + Scc1': Smc3-StrepII/Smc1-6×His + Scc1-6×His
'Cys dimer + Scc1': Smc3(S1043C)-StrepII/Smc1-6×His + Scc1-6×His
'WT hinge': Smc3-StrepII/Smc1-6×His
'Cys hinge': Smc3(E570C)-StrepII/Smc1(L639C)-6×His]

stoichiometric ratio. This not being the case points to the co-elution being non-specific. Physiological interaction between the components of the cohesin trimer may well require the presence of Scc3, ATP, or another co-factor for proper assembly *in vivo* and these criteria may have not been met in the reconstitution attempt presented in this thesis.

(iii) Another plausible explanation for lack of detectable, functional assembly of the 'trimer' is that the purification method used was unsuitable for isolation of correctly-folded Scc1 despite the presence of an SMC dimer 'poised' for binding. As discussed earlier, purification of recombinant, eukaryotic proteins from bacterial expression systems can pose problems of solubility and folding due to the absence of chaperones that 'guide' a protein's hydrophobic residues into their physiological conformation. It is possible, for example, that the soluble Scc1 material was misfolded or forming aggregates to begin with or that this was induced by the multiple rounds of purification. This is less likely to be the case for the SMC dimer, as functional assembly of the hinge domains was demonstrated in Figure 13. One option to improve Scc1's solubility would be to use an affinity tag that boosts recombinant protein solubility (e.g. maltose-binding protein tag, MBP tag) or switch to an eukaryotic expression system such as insect cells (O'Shaughnessy and Doyle, 2011). Alternatively, the quality of protein obtained during co-expression of Scc1 with other core cohesin subunits should be investigated further and pursued if the results are promising.

(iv) It is possible that the purification strategy was not the reason behind the lack of 'trimer' cross-linking but the cysteine pair used is not cross-linkable *in*

vitro. While Smc3(S1043C) and Scc1(C56) have been shown to cross-link *in vivo*, it is not a given that this will happen in an artificial system with only a few purified components. The reason for this idea is the following: in the crystal structure of the Smc3 head domain with the N-terminal part of Scc1, Smc3's S1043 is located too far for chemical cross-linking to occur but the Smc3(S1043C)/Scc1(C56) interaction can be reproducibly captured by *in vivo* cross-linking (Gligoris et al., 2014). This implies that this interaction is dependent on flexible sections of either or both proteins periodically coming together. There may, for example, be a conformational change that needs to occur before the two cysteines are close enough to be bridged by a cross-linker. In an *in vivo* situation, there is presumably always a population of cohesin that has undergone this change and a cross-linking signal will be detected. To investigate this possibility, a pair of cysteine residues designed on the basis of the crystal structure to be a theoretically and practically cross-linkable distance apart should be used.

Co-purification with Scc3 boosts Scc1's solubility almost two-fold, indicating that tetramer reconstitution may be a viable option

One reason for the 'trimer' reconstitution not being functionally assembled was related to the misfolding and/or poor solubility of Scc1. Co-expression of Scc1-6×His with Smc1-FLAG resulted in production of Scc1 that was effectively completely soluble at 80%, as determined by quantification of 'total' versus 'soluble' Western blot lanes in Figure 9D. In theory, co-expression of Scc1 with

another binding partner, like Scc3, should have a similar effect and increase the probability of purifying a functional version of Scc1. Since the SMC dimer had already been shown to assemble functionally and co-purifying Scc1 and Scc3 was a feasible option, then cohesin tetramer assembly could be achieved in a simple, two-step manner.

Using non-codon-optimised yeast *SCC3*, a 'bi-cistronic' Scc3-StrepII/Scc1-6×His construct was cloned and introduced into a bacterial expression strain (Figure 19A). IPTG-mediated induction of expression resulted in production of a modest amount of both proteins of interest, detectable by Western blotting but not Coomassie-staining in the WCE (Figure 19B). 4 litres of induced bacterial culture were used for His-tag affinity purification, which generated a large relatively broad absorbance peak (Figure 19C). Comparison of Western blot quantification values obtained from the 'total' and 'soluble' lanes revealed that 58% of the Scc1 expressed was soluble, an almost two-fold improvement on expression of Scc1 alone (Figure 19D). Scc3-StrepII was detectable only in fractions in the sharpest section of the absorbance curve while Scc1-6×His was present throughout the wide peak. As before, Coomassie-staining of the total protein established that the elution contained non-specific contaminants but these could presumably be removed by ion exchange chromatography.

Unfortunately, due to time constraints, cohesin tetramer reconstitution using the co-purified Scc3/Scc1 sample and the previously obtained SMC dimer could not be attempted. The next step would be to increase the purity of the Scc3/Scc1 elution by anion exchange chromatography, gel filtrate it in the presence of the

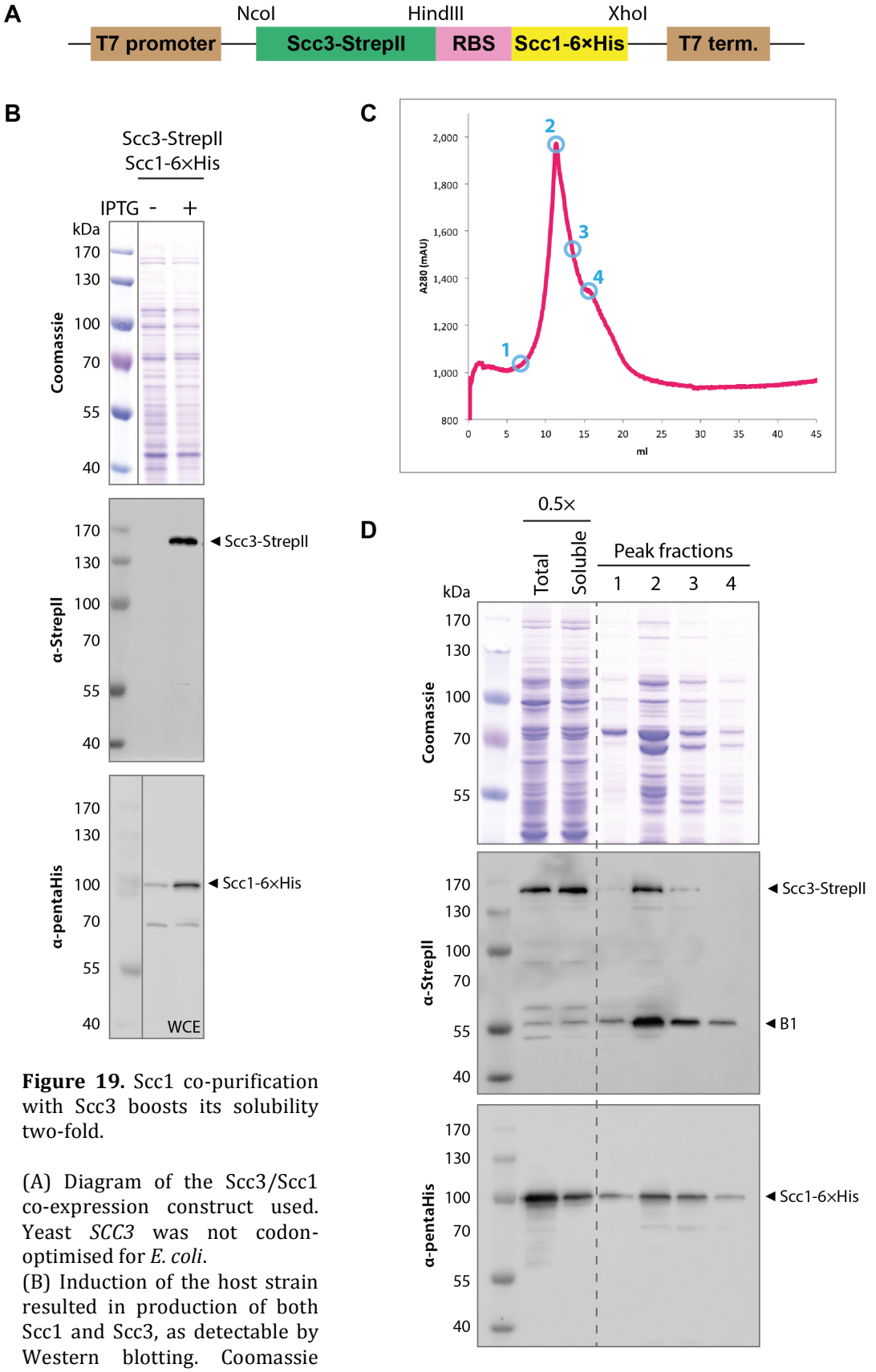

Figure 19. Scc1 co-purification with Scc3 boosts its solubility two-fold.

(A) Diagram of the Scc3/Scc1 co-expression construct used. Yeast *SCC3* was not codon-optimised for *E. coli*.
(B) Induction of the host strain resulted in production of both Scc1 and Scc3, as detectable by Western blotting. Coomassie staining of the WCE demonstrated that both proteins were produced at relatively low levels.
(C) His-tag purification from a 4 litre culture generated a large, broad absorbance peak.
(D) The soluble pool of Scc1 was almost 60% and this material co-purified with Scc3. The latter was most abundant in the sharpest section of the absorbance spectrum, while Scc1 was present throughout the observed peak. [**Note:** The band marked 'B1' was assumed to be a cleavage product of the Scc3-StrepII protein.]

SMC dimer, and then verify tetramer assembly. Verification of correct complex formation should be done by co-IP followed by attempting to chemically cross-link Scc1 to the SMCs while keeping in mind the technical concerns raised previously (i.e. better positive control, more material, a different cysteine pair in the Smc3/Scc1 interface). What is promising is that at least part of the soluble Scc1 pool must be folding physiologically to enable its co-elution with Scc3. This should be pursued further, as it is a good indicator that cohesin tetramer reconstitution may be a viable strategy.

Discussion

The expression and purification system described in this thesis to express *S. cerevisiae* cohesin subunits using 'bi-cistronic', codon-optimised constructs in a bacterial expression system holds viable potential. As illustrated by the successful purification of the Smc1/Smc3 heterodimer, whose functional interaction was verified by site-specific cross-linking at the hinge interface, high yields of good-quality material could be reliably purified with relative ease and speed. The His-tag purified material contained a large proportion of protein contaminants that required removal by size exclusion chromatography; this issue could have been avoided by using the more specific StrepII-tag. In future, the least-expressed protein in a 'poly-cistronic' construct should always be tagged with StrepII (e.g. Smc1-StrepII in the SMC dimer construct) to reduce the need for a second purification step.

Although not attempted in this thesis due to time constraints, the functionally assembled SMC dimer in itself could be used to further our understanding of the dynamics of cohesin's DNA 'entry gate.' The currently established technique to measure the association between Smc1 and Smc3 is a competition assay using purified SMC hinge domains where a mixture of two SMC hinges is supplemented with an excess of one uniquely labelled SMC hinge fragment. During a time-course experiment, displacement of one of the hinge fragments from the original mix by the competitor hinge is measured by Western blotting (Kurze et al., 2011). Since truncated versions of the SMCs are used, this approach can only reproduce properties affecting hinge dimerisation intrinsic to the hinge domains themselves. What it does not consider is the potential influence that the rest of an SMC's sequence has on its hinge domain and vice versa. One could use full-length SMCs purified using the approach described in this thesis to create a more physiological version of the Smc1/Smc3 competition assay described above.

Although functional assembly of the reconstituted 'trimer' using the purified yeast SMC dimer and Scc1 could not be verified by chemical cross-linking, this could be explained by several factors as discussed earlier. Most likely, expression and purification of Scc1 alone was detrimental to the protein's folding. Co-expression of Scc1 with either Smc1 or Scc3 was shown to greatly increase the amount of soluble Scc1 that co-eluted with the protein in question. This co-expression approach should be pursued further for Scc1, as co-purification with specific binding partners most likely results from a correctly folded interaction interface and, presumably, a functionally assembled protein. The most promising

avenue to pursue is the co-expression of Scc1 with Scc3 to complement the functionally assembled SMC dimer.

CHAPTER IV

Summary and future directions

CHAPTER IV: Summary and future directions

Summary

Cohesin, a protein complex composed of four subunits arranged to form a large proteinaceous ring, is an essential player involved in timely and accurate segregation of genetic material during mitosis. In more complex eukaryotes, stabilisation of cohesin on DNA depends on the presence of sororin, a small functionally conserved protein that is thought to counteract the role of Wapl, which in turn is key to releasing cohesin from chromatin in a cleavage-independent manner.

The work described in the first half of this thesis aimed to address the discrepancy of *S. cerevisiae* not possessing a sororin orthologue despite its conserved function in other organisms. An imaging-based screen of the yeast genome was performed, streamlined such that only candidates with mRNAs regulated similarly to cohesin subunits' were considered. The screen did not highlight any feasible candidates for 'yeast sororin' or any other novel cohesin regulators. A more inclusive, mass spectroscopy-based approach may be able to identify novel cohesin-associated regulatory proteins, but this work was not undertaken here.

The work described in the latter half of this thesis concerned the attempt to purify the *S. cerevisiae* cohesin complex using a bacterial expression system. This was done to establish a reliable, efficient method for recombinant cohesin complex isolation for future use in biochemical assays to address questions about cohesin loading and release that are difficult to answer *in vivo*. The Smc1/Smc3 heterodimer was successfully purified and characterised to show that its hinge dimerisation domain was functionally assembled. Following expression of Scc1 alone, cohesin 'trimer' reconstitution could not be verified by chemical cross-linking. That said, an alternative approach of co-purifying Scc1 with Scc3 was demonstrated and should be used to attempt cohesin 'tetramer' reconstitution in the near future.

Future directions

Screen for yeast sororin orthologue

Not being able to identify a sororin orthologue in *S. cerevisiae*, or explain how its requirement is bypassed should it not exist, highlights the limitations of our understanding of sister chromatid cohesion. As discussed previously, arguably the most feasible method of identifying a yeast sororin orthologue would be to perform a more inclusive and less subjective screen. Such an approach should be based on current knowledge of sororin's function and include as few assumptions as possible. Thus, the two most viable options would be to analyse the whole yeast proteome either for novel proteins whose association with

cohesin depends on Smc3 acetylation or for additional targets of APC/C^{Cdh1} in a modified version of the approach that originally identified *X. laevis* sororin. That said, perhaps a more sensible tactic altogether would be to first try and characterise sororin's function in greater molecular detail using an organism with a confirmed sororin orthologue. This way any further screening of the yeast genome could be more directed towards a specific phenotype.

Cohesin complex reconstitution for use in biochemical assays

Setting up an *in vitro* assay capable of reconstituting the cohesin loading and releasing reactions using budding yeast components remains an important goal to the advancement of our understanding of the regulation of sister chromatid cohesion. Given the success of the expression system presented in this thesis with the SMC dimer, it would be wise to pursue the promising approach of co-purifying of Scc1 with Scc3. Alternatively, given that co-expression of Scc1 with other binding partners boosts its solubility, perhaps expressing all four subunits of the core cohesin complex in one 'poly-cistronic' vector should be attempted. If either of these strategies were to result in a functionally assembled tetramer, cryo-EM studies of the complex's architecture could begin. Meanwhile, efforts would also need to shift to isolating cohesin's regulatory subunits, most crucially Scc2, in an attempt at reconstituting the cohesin loading reaction.

CHAPTER V

Materials and methods

CHAPTER V: Materials and methods

Yeast GFP clone collection

The entire library of available yeast protein GFP fusions (Huh et al., 2003) was obtained from Life Technologies. Strains were recovered from supplied stocks by patching on –histidine drop-out plates.

Live cell imaging

Yeast cells were grown at 25°C at 200 rpm in a 20 ml YPD culture from OD_{600} 0.1 to log phase (~4 hours) from an overnight starter culture. Cells were collected by centrifugation at 2,500 *g* for 3 min, the liquid decanted and the cells resuspended in the small volume of media remaining at the bottom of the tube. 1 µl of the live cell suspension was spread out thinly on a microscopy slide and visualised using the DeltaVision Elite (DeltaVision) system with a 100×/1.40 NA objective. Image analysis was performed in Fiji (Schindelin et al., 2012).

Plasmid construction for protein expression

S. cerevisiae cohesin genes *SMC1*, *SMC3*, and *SCC1* were codon-optimised for *E. coli* and synthesised by GenScript. They were then subcloned into the pET28a expression vector (Novagen) with either a 6×His, StrepII, or FLAG (sequences in table below) carboxy-terminal tag (unless stated otherwise). An extra alanine was also added just after the start codon to avoid N-terminal degradation of the

recombinant protein (Varshavsky, 1997). Transcription was regulated by the IPTG-inducible T7lac promoter driving the expression of either a single gene or two genes in tandem separated by a short ribosome binding sequence (RBS; DNA sequence: GTTTAACTTTAAGAAGGAGATATAAAT) to create a 'bi-cistronic' construct. Oligonucleotides (oligos) used for cloning, sequencing, and site-directed mutagenesis are listed in Appendix 1. Plasmids generated during the course of this project are listed in Appendix 2.

Affinity tag	Amino acid sequence	DNA sequence
6×His	HHHHHH	CATCACCACCACCACCAC
StrepII	WSHPQFEK	TGGTCACATCCGCAGTTTGAGAAA
FLAG	DYDDDK	GATTATGATGACGATGACAAA

(i) PCR amplification

PCR amplification of desired genes was performed using Q5 High-Fidelity DNA Polymerase (New England Biolabs) and DNA oligonucleotides (Life Technologies) containing appropriate tags and restriction enzyme sites in the following reaction mix:

Component	Volume added (µl)	
	1 × 50µl reaction	4.5 × 50µl reactions
ddH_2O	34.5	155.25
5× Q5 reaction buffer	10	45
Forward oligo (20 µM)	1	4.5
Reverse oligo (20 µM)	1	4.5
dNTPs (10 mM each; Roche)	0.5	2.25
DMSO	1.5	6.75
DNA template (~100 ng/ul)	1	4.5
Q5 polymerase	0.5	2.25

The PCR reaction mix was then cycled in a Tetrad 2 Peltier Thermal Cycler (Bio-Rad) under the following conditions:

Step	Temperature	Duration	
Initial denaturation	98°C	5 min	
Denaturation	98°C	10 sec	30 cycles
Annealing*	65°C	30 sec	
Extension	72°C	30 sec per kb of expected product	
Final extension	72°C	5 minutes	
Hold	15°C	∞	

*Primer annealing was performed either at 65°C or at the specific melting temperature (T_m) for a given pair of primers, as calculated using A Plasmid Editor (ApE) software.

5 µl of the PCR reaction plus 1 µl of 6× DNA Loading Dye (Thermo Scientific) were then ran on an agarose gel (1% w/v in TAE with 0.2 µg/ml ethidium bromide) to confirm the presence and size of the product. Nucleotides, salts, and other contaminants were then removed from the reaction using the QIAquick PCR Purification Kit (Qiagen) prior to restriction enzyme digestion.

(ii) Restriction enzyme digestion

Restriction enzyme digestion was performed using NcoI-HF, HindIII-HF, and XhoI restriction enzymes (New England Biolabs) overnight at 37°C, in a reaction mix detailed in the manufacturer's instructions. Restriction digests were cleaned up using the QIAquick PCR Purification Kit (Qiagen) as detailed before or, in the case of plasmids where large inserts were removed, using the QIAquick Gel

Extraction Kit (Qiagen). The DNA concentration was then quantified based on its absorbance at 260 nm, as measured by a NanoDrop-1000 spectrophotometer (Thermo Scientific).

(iii) Ligation of inserts into pET28a backbone

The following formula was used to calculate the amount of insert required to achieve an insert:vector ratio of 3:1:

$$\text{ng of insert required} = \frac{(\text{ng of vector}) \times (\text{size}_{\text{insert}} \text{ in kb})}{\text{size}_{\text{vector}} \text{ in kb}} \times \text{insert:vector ratio}$$

Using the TaKaRa DNA Ligation Kit (Mobitec) as follows, the digested insert was then ligated into the pET28a backbone, which had been linearised with the appropriate enzyme(s):

Component	Amount added
Linearised vector backbone	50 ng
Digested insert	as calculated above
TaKaRa solution I	volume equal to volume of (vector + insert)
ddH_2O	make up to 10 µl

The ligation mix was incubated at 16°C for 30 min and then transformed into XL1-Blue competent cells.

(iv) Bacterial transformation

To amplify plasmid DNA, ligation reactions, site-directed mutagenesis reactions, and purified plasmids were transformed into XL1-Blue chemically competent

E. coli cells. 10 µl of a ligation mix, 5 µl of a mutagenesis reaction, or 1 µl of a pure plasmid was incubated with 50 µl of cells for 20 min on ice. Cells were then heat shocked at 42°C for 45 sec, followed by 2 min on ice. 1 ml of 2×TY medium was then added to the transformation mix and the samples incubated for 1 hour at 37°C, 220 rpm. For ligations and mutagenesis, the cells were pelleted (16,000 *g*, 30 sec), most of the supernatant tipped out such that the pellet could be resuspended in approximately 100 µl of liquid, and spread on LB agar plates containing 50 mg/ml kanamycin. For purified plasmid transformations, 50 µl of the unpelleted transformation mix was plated. Plates were incubated at 37°C for 16-20 hours.

(v) Plasmid recovery and verification of ligation products

Plasmids were recovered from a 5 ml overnight culture in 2×TY medium containing antibiotics using a QIAprep Spin Miniprep Kit (Qiagen). Correct fragment insertion was verified by a small-scale restriction digest (5 µl of the Miniprep elution) followed by agarose gel electrophoresis. Clones with correct digestion patterns were diluted to 100 ng/µl in 5 µl and sent for Sanger sequencing (Source BioScience) with gene-specific primers (listed in Appendix 1) or those in the T7 promoter (T7F: TAATACGACTCACTATAGGG) and T7 terminator (T7R: GCTAGTTATTGCTCAGCGG) regions.

(vi) Site-directed mutagenesis

Oligos for site-directed mutagenesis (SDM) were designed to be 29 base pairs (bp) in length. The codon to be mutagenised was placed 5 bp from the end of the

oligo such that the overlap between the forward and reverse oligos was 8 bp with 21 bp of homology to the template plasmid on either side, ensuring robust binding. The mutagenesis mix was set up like a standard amplification PCR (see above), except that oligos were used at 100 µM and the extension time was 3 min 45 sec. Each 50 µl PCR reaction was incubated overnight at 37°C with 1 µl of DpnI (New England Biolabs) to digest away the methylated template DNA. 5 µl of the reaction was then transformed into 50 µl of competent XL1-Blue cells, the plasmids recovered, and the presence of the desired mutation(s) verified by sequencing, as described previously.

SDS-PAGE and protein detection

All protein samples were run on 8% polyacrylamide gels at 150V for 1 hour 20 min. For Coomassie staining, gels were microwaved for 2×1 min at full power in distilled water, the water replaced with Imperial Protein Stain (Thermo Scientific), microwaved for a further 2×30 sec, and incubated for at least 10 min with agitation. The stain was then discarded, the gels microwaved for a final 1 min in distilled water, and incubated up to overnight with shaking before visualisation. For Western blotting, proteins were transferred onto a 0.2 µm PVDF membrane using the Trans-Blot Turbo Transfer System (Bio-Rad) for 10 min at 1.3 A (up to 25 V) per gel. The membrane was blocked in appropriate blocking solution (see table below) for at least 30 min at RT with shaking. It was then incubated with the required primary antibody either for 1.5 hours at RT or overnight at 4°C, followed by an incubation with the secondary antibody for 45 min at RT. All blots were washed with phosphate buffered saline + 0.05% Tween

20 (PBS-T; Sigma) for approximately 5 min between antibody incubation steps. Blots were developed with Immobilon Western Chemiluminescent HRP Substrate (Merck Millipore) using an ODYSSEY Fc Imaging System (LI-COR Biosciences). Western blots were quantified using Fiji software (Schindelin et al., 2012).

(i) Primary antibodies

Epitope	Supplier	Blocking solution	Working dilution	Origin
Penta-His	Qiagen	3% w/v BSA (Sigma) in PBS-T	1:1,000 in 3% w/v BSA (Sigma) in PBS-T	Mouse
StrepII-tag	Merck Millipore		1:1,000 in PBS-T	
FLAG-tag	Sigma	5% w/v skimmed milk (Sigma) in PBS-T	1:1,000 in 5% w/v skimmed milk (Sigma) in PBS-T	Rabbit

(ii) Secondary antibodies

Target species	Supplier	Working dilution
Mouse	GE Healthcare	1:10,000 in 5% w/v skimmed milk (Sigma) in PBS-T
Rabbit		

Protein expression tests

For expression tests, 500 µl of an overnight bacterial culture was pelleted at 16,000 *g*, 30 sec. 150 µl of 2× SDS loading buffer was added to the samples, which were subsequently denatured for 2 min at 95°C, sonicated for 10 sec at 20% output (Vibra-Cell, Sonics) to reduce the viscosity of the sample, and spun down for 5 min at 16,000 *g*. 5 µl of each sample was run on an 8%

polyacrylamide gel (as described above) and the presence of the protein(s) of interest determined by both Coomassie staining and Western blotting.

Large-scale protein expression

The pET28a vector in BL21 DE3 *E. coli* cells was used for all protein expression. Strains generated during this project are listed in Appendix 2. Cultures from single bacterial colonies were inoculated into liquid 2×TY medium with appropriate antibiotics (see table below) and grown overnight at 37°C, 220 rpm. Cultures were diluted to OD_{600} 0.1 in fresh media with antibiotics and grown for 2 hours before shifting the temperature of the incubator to 16°C. Approximately 1.5 hours later, once the incubator had cooled to 20°C or below, IPTG was added to a final concentration of 0.4 mM and the cultures incubated overnight at 16°C, 220 rpm. Cells were harvested the next morning using a JLA-8.1000 rotor in an Avanti J-26S XP centrifuge (Beckman Coulter, Inc.) for 15 min, 4000 *g*, 4°C. Cell pellets were either lysed immediately or flash frozen in liquid nitrogen for storage at -80°C.

Antibiotic	Supplier	1000× stock concentration (w/v)
Ampicillin	Sigma	50 mg/ml in sterile H_2O
Kanamycin	Sigma	50 mg/ml in sterile H_2O
Chloramphenicol	Fluka BioChemika	34 mg/ml in 100% EtOH

Affinity tag purification

Pellets from up to 8 litres of induced *E. coli* cultures were resuspended in 3 times the pellet volume of appropriate lysis buffer (see table below), which also contained one complete EDTA-free protease inhibitor cocktail tablet (Roche Diagnostics) per 50 ml solution:

Affinity tag	Lysis and wash buffer	Composition	Eluent
His-tag	Buffer H	50mM Tris-HCl (pH 7.5) 300mM NaCl 10% glycerol (v/v)	Imidazole
StrepII-tag	Buffer S	100mM Tris-HCl (pH 8.0) 150mM NaCl 1mM EDTA 10% glycerol (v/v)	d-Desthiobiotin

The cells were lysed using a French pressure cell press (Constant Systems Ltd.) at 20 kpsi, 4°C. Lysates were supplemented with PMSF (from a 100 mM stock solution in 100% EtOH) to a final concentration of 1 mM prior to sonication (Vibra-Cell, Sonics) on ice for 2×30 sec at 80% output per 50 ml lysate to shear the DNA and reduce the viscosity of the sample. Lysates were then centrifuged in a JA-25.50 rotor in an Avanti J-26S XP centrifuge (Beckman Coulter, Inc.) at 50,000 *g* for 90 min at 4°C. Using an ÄKTApurifier 10 system (GE Healthcare), cleared lysates were injected at a flow of 0.5 ml/min onto a HisTrap HP or StrepTrap HP 5 ml pre-packed affinity column (GE Healthcare), which had been pre-equilibrated in buffer H or S, respectively. Flow-through was collected and the column washed with the appropriate buffer until the UV 280 nm reading reached a plateau. Proteins were eluted using a gradient elution from 0% to 100% of suitable elution buffer (EB; see below) spread across 8 column volumes

(CVs) with an additional 5 CV wash at 100% EB.

Affinity tag	Elution buffer (EB)	Composition
His-tag	Buffer EB-H	50mM Tris-HCl (pH 7.5) 300mM NaCl 10% glycerol (v/v) **500 mM imidazole**
StrepII-tag	Buffer EB-S	100mM Tris-HCl (pH 8.0) 150mM NaCl 1mM EDTA 10% glycerol (v/v) **10 mM d-Desthiobiotin**

Samples from peak fractions were taken for protein detection while the remaining volume was transferred into microfuge tubes and either processed immediately or flash frozen in liquid nitrogen for storage at -80°C.

Ion exchange chromatography

After analysis by Coomassie staining and Western blotting, fractions containing the protein(s) of interest following affinity tag purification were diluted 10× in buffer S. This was then injected at 1 ml/min onto a 1 ml HiTrapQ HP anion exchange column (GE Healthcare), which had first been regenerated with buffer S_{1000} (see table below) and then equilibrated in buffer S. After washing with buffer S until the UV 280 nm reading stabilised, proteins were eluted using a gradient of increasing concentration of buffer S_{500} across 20 CVs, followed by a final wash of 10 CVs of 100% buffer S_{500}.

Buffer	Purpose	Composition
Buffer S_{500}	Protein elution	100mM Tris-HCl (pH 8.0) **500 mM NaCl** 1mM EDTA 10% glycerol (v/v)
Buffer S_{1000}	Column regeneration	100mM Tris-HCl (pH 8.0) **1 M NaCl** 1mM EDTA 10% glycerol (v/v)

Gel filtration

Using an ÄKTApurifier 10 system (GE Healthcare), a 300 mm Superose 6 10/300 GL size exclusion column (GE Healthcare) was equilibrated with StrepII-tag LB (see above) until the conductivity reached a plateau. Gel filtration standards (Bio-Rad) were ran through the column at a flow rate of 0.1-0.3 ml/min. 100 µl fractions were collected in 96-well plates over 30 ml (1.2× CV). Up to 500 µl of protein sample was then injected manually into the column and fractions collected as before. One volume of 2× SDS loading buffer was added to samples from each fraction to be analysed and denatured for 2 min at 95°C prior to gel electrophoresis.

Determination of protein concentration

Protein concentration was measured based on the absorbance at 280 nm, as determined by a NanoDrop-1000 spectrophotometer (Thermo Scientific). To increase the accuracy of the readings, the calculated extinction coefficient and molecular weight (Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., 2005) of the protein being quantified was entered into the NanoDrop software.

In vitro cross-linking

For SMC hinge cross-linking, 100 µl of purified material at approximately 5 mg/ml in buffer S was used per reaction. For 'trimer' assembly cross-linking, one subcomplex was diluted in buffer S to match its concentration to that of the other subcomplex. Cross-linker stock solutions were made up fresh as follows in dimethyl sulphoxide (DMSO):

Cross-linker	Mass used per 200 µl DMSO	Stock concentration	Working concentration
BMOE	1.10 mg	25 mM	2 mM
dBBr	0.35 mg	5 mM	0.4 mM

8 µl of cross-linker stock solution or DMSO (negative control) were added to 100 µl of protein sample and the mix incubated on ice for 10 min under a light-proof cover. The reaction was terminated by addition of 25 µl of 4× LDS sample buffer (Life Technologies) and incubation at 95°C for 5 min. Up to 30 µl of each sample were then loaded onto a 3-8% Tris-Acetate 1.0 mm protein gel (Life Technologies) and ran at 45 mA per gel for 1 hour 10 min. Proteins were transferred onto a 0.45 µm Immobilon-P PVDF membrane (Merck Millipore) using a semi-dry blotter (Hoefer Inc.) at 144 mA for 2 hours (one or two membranes simultaneously) prior to detection by Western blotting as described previously.

CHAPTER VI

References

CHAPTER VI: References

Anderson, D.E., Losada, A., Erickson, H.P., and Hirano, T. (2002). Condensin and cohesin display different arm conformations with characteristic hinge angles. J. Cell Biol. *156*, 419–424.

Arumugam, P., Gruber, S., Tanaka, K., Haering, C.H., Mechtler, K., and Nasmyth, K. (2003). ATP hydrolysis is required for cohesin's association with chromosomes. Curr. Biol. *13*, 1941–1953.

Ben-Shahar, T.R., Heeger, S., Lehane, C., East, P., Flynn, H., Skehel, M., Uhlmann, F., Rolef Ben-Shahar, T., Heeger, S., Lehane, C., et al. (2008). Eco1-Dependent Cohesin Acetylation During Establishment of Sister Chromatid Cohesion. Science *321*, 563–566.

Blat, Y., and Kleckner, N. (1999). Cohesins Bind to Preferential Sites along Yeast Chromosome III, with Differential Regulation along Arms versus the Centric Region. Cell *98*, 249–259.

Bürmann, F., Shin, H.-C., Basquin, J., Soh, Y.-M., Giménez-Oya, V., Kim, Y.-G., Oh, B.-H., and Gruber, S. (2013). An asymmetric SMC-kleisin bridge in prokaryotic condensin. Nat. Struct. Mol. Biol. *20*, 371–379.

Chan, K.-L., Roig, M.B., Hu, B., Beckouët, F., Metson, J., and Nasmyth, K. (2012). Cohesin's DNA exit gate is distinct from its entrance gate and is regulated by acetylation. Cell *150*, 961–974.

Chan, K.-L., Gligoris, T., Upcher, W., Kato, Y., Shirahige, K., Nasmyth, K., and Beckouët, F. (2013). Pds5 promotes and protects cohesin acetylation. Proc. Natl. Acad. Sci. U. S. A. *110*, 13020–13025.

Chandler, D. (2005). Interfaces and the driving force of hydrophobic assembly. Nature *437*, 640–647.

Ciosk, R., Zachariae, W., Michaelis, C., Shevchenko, A., Mann, M., and Nasmyth, K. (1998). An ESP1/PDS1 Complex Regulates Loss of Sister Chromatid Cohesion at the Metaphase to Anaphase Transition in Yeast. Cell *93*, 1067–1076.

Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A., Shevchenko, A., and Nasmyth, K. (2000). Cohesin's Binding to Chromosomes Depends on a Separate Complex Consisting of Scc2 and Scc4 Proteins. Mol. Cell *5*, 243–254.

Eichinger, C.S., Kurze, A., Oliveira, R.A., and Nasmyth, K. (2013). Disengaging the Smc3/kleisin interface releases cohesin from Drosophila chromosomes during interphase and mitosis. EMBO J. *32*, 656–665.

Eith, C., Kolb, M., and Seubert, A. (2001). Practical Ion Chromatography: An Introduction (Herisau, Switzerland: Metrohm. Ltd).

Eser, P., Demel, C., Maier, K.C., Schwalb, B., Pirkl, N., Martin, D.E., Cramer, P., and Tresch, A. (2014). Periodic mRNA synthesis and degradation co-operate during cell cycle gene expression. Mol. Syst. Biol. *10*, 717.

Farcas, A.-M., Uluocak, P., Helmhart, W., and Nasmyth, K. (2011). Cohesin's Concatenation of Sister DNAs Maintains Their Intertwining. Mol. Cell *44*, 97–107.

Ferrer, M., Chernikova, T.N., Yakimov, M.M., Golyshin, P.N., and Timmis, K.N. (2003). Chaperonins govern growth of Escherichia coli at low temperatures. Nat. Biotechnol. *21*, 1266–1267.

Feytout, A., Vaur, S., Genier, S., Vazquez, S., and Javerzat, J.-P. (2011). Psm3 acetylation on conserved lysine residues is dispensable for viability in fission yeast but contributes to Eso1-mediated sister chromatid cohesion by antagonizing Wpl1. Mol. Cell. Biol. *31*, 1771–1786.

Foley, E.A., and Kapoor, T.M. (2013). Microtubule attachment and spindle assembly checkpoint signalling at the kinetochore. Nat Rev Mol Cell Biol *14*, 25–37.

Gandhi, R., Gillespie, P.J., and Hirano, T. (2006). Human Wapl Is a Cohesin-Binding Protein that Promotes Sister-Chromatid Resolution in Mitotic Prophase. Curr. Biol. *16*, 2406–2417.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., B.A. (2005). The Proteomics Protocols Handbook. In Protein Identification and Analysis Tools on the ExPASy Server, J.M. Walker, ed. (Humana Press), pp. 571–607.

Gerlich, D., Koch, B., Dupeux, F., Peters, J.-M.M., and Ellenberg, J. (2006). Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication. Curr Biol *16*, 1571–1578.

Gligoris, T.G., Scheinost, J.C., Burmann, F., Petela, N., Chan, K.-L., Uluocak, P., Beckouet, F., Gruber, S., Nasmyth, K., and Lowe, J. (2014). Closing the cohesin ring: Structure and function of its Smc3-kleisin interface. Science *346*, 963–967.

Glynn, E.F., Megee, P.C., Yu, H.-G., Mistrot, C., Unal, E., Koshland, D.E., DeRisi, J.L., and Gerton, J.L. (2004). Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae. PLoS Biol. *2*, E259.

Grantham, R., Gautier, C., Gouy, M., Mercier, R., and Pavé, A. (1980). Codon catalog usage and the genome hypothesis. Nucleic Acids Res. *8*, r49–r62.

Gruber, S., Haering, C.H., and Nasmyth, K. (2003). Chromosomal Cohesin Forms a Ring. Cell *112*, 765–777.

Gruber, S., Arumugam, P., Katou, Y., Kuglitsch, D., Helmhart, W., Shirahige, K., and Nasmyth, K. (2006). Evidence that Loading of Cohesin Onto Chromosomes Involves Opening of Its SMC Hinge. Cell *127*, 523–537.

Guacci, V., Koshland, D., and Strunnikov, A. (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S. cerevisiae. Cell *91*, 47–57.

Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. Trends Biotechnol. *22*, 346–353.

Haering, C.H., Löwe, J., Hochwagen, A., and Nasmyth, K. (2002). Molecular Architecture of SMC Proteins and the Yeast Cohesin Complex. Mol. Cell *9*, 773–788.

Haering, C.H., Schoffnegger, D., Nishino, T., Helmhart, W., Nasmyth, K., and Löwe, J. (2004). Structure and Stability of Cohesin's Smc1-Kleisin Interaction. Mol. Cell *15*, 951–964.

Haering, C.H., Farcas, A.-M., Arumugam, P., Metson, J., and Nasmyth, K. (2008). The cohesin ring concatenates sister DNA molecules. Nature *454*, 297–301.

Hartman, T., Stead, K., Koshland, D., and Guacci, V. (2000). Pds5p is an essential chromosomal protein required for both sister chromatid cohesion and condensation in Saccharomyces cerevisiae. J. Cell Biol. *151*, 613–626.

Heidinger-Pauli, J.M., Onn, I., and Koshland, D. (2010). Genetic evidence that the acetylation of the Smc3p subunit of cohesin modulates its ATP-bound state to promote cohesion establishment in Saccharomyces cerevisiae. Genetics *185*, 1249–1256.

Hirano, M., and Hirano, T. (2002). Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA. EMBO J. *21*, 5733–5744.

Hoenger, A. (2014). High-resolution cryo-electron microscopy on macromolecular complexes and cell organelles. Protoplasma *251*, 417–427.

Hu, B., Itoh, T., Mishra, A., Katoh, Y., Chan, K.-L., Upcher, W., Godlee, C., Roig, M.B., Shirahige, K., and Nasmyth, K. (2011). ATP Hydrolysis Is Required for Relocating Cohesin from Sites Occupied by Its Scc2/4 Loading Complex. Curr. Biol. *21*, 12–24.

Huh, W.K., Falvo, J. V, Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686–691.

Ikemura, T. (1981). Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the E. coli translational system. J. Mol. Biol. *151*, 389–409.

Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. *2*, 13–34.

Ivanov, D., Schleiffer, A., Eisenhaber, F., Mechtler, K., Haering, C.H., and Nasmyth, K. (2002). Eco1 Is a Novel Acetyltransferase that Can Acetylate Proteins Involved in Cohesion.

Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. *292*, 195–202.

Kueng, S., Hegemann, B., Peters, B.H., Lipp, J.J., Schleiffer, A., Mechtler, K., and Peters, J.-M. (2006). Wapl Controls the Dynamic Association of Cohesin with Chromatin. Cell *127*, 955–967.

Kurze, A., Michie, K.A., Dixon, S.E., Mishra, A., Itoh, T., Khalid, S., Strmecki, L., Shirahige, K., Haering, C.H., Lowe, J., et al. (2011). A positively charged channel within the Smc1/Smc3 hinge required for sister chromatid cohesion. EMBO J *30*, 364–378.

Lafont, A.L., Song, J., and Rankin, S. (2010). Sororin cooperates with the acetyltransferase Eco2 to ensure DNA replication-dependent sister chromatid cohesion. Proc. Natl. Acad. Sci. U. S. A. *107*, 20364–20369.

Laloraya, S., Guacci, V., and Koshland, D. (2000). Chromosomal addresses of the cohesin component Mcd1p. J. Cell Biol. *151*, 1047–1056.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947–2948.

Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G.P., Itoh, T., Watanabe, Y., Shirahige, K., and Uhlmann, F. (2004). Cohesin relocation from sites of chromosomal loading to places of convergent transcription. Nature *430*, 573–578.

Liu, D., and Lampson, M.A. (2009). Regulation of kinetochore-microtubule attachments by Aurora B kinase. Biochem. Soc. Trans. *37*, 976–980.

Lopez-Serra, L., Lengronne, A., Borges, V., Kelly, G., and Uhlmann, F. (2013). Budding Yeast Wapl Controls Sister Chromatid Cohesion Maintenance and Chromosome Condensation. Curr. Biol. *23*, 64–69.

Megee, P.C., Mistrot, C., Guacci, V., and Koshland, D. (1999). The Centromeric Sister Chromatid Cohesion Site Directs Mcd1p Binding to Adjacent Sequences. Mol. Cell *4*, 445–450.

Melby, T.E., Ciampaglio, C.N., Briscoe, G., and Erickson, H.P. (1998). The Symmetrical Structure of Structural Maintenance of Chromosomes (SMC) and MukB Proteins: Long, Antiparallel Coiled Coils, Folded at a Flexible Hinge. J. Cell Biol. *142*, 1595–1604.

Michaelis, C., Ciosk, R., and Nasmyth, K. (1997). Cohesins: Chromosomal Proteins that Prevent Premature Separation of Sister Chromatids. Cell *91*, 35–45.

Miller, C., Schwalb, B., Maier, K., Schulz, D., Dumcke, S., Zacher, B., Mayer, A., Sydow, J., Marcinowski, L., Dolken, L., et al. (2011). Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. Mol Syst Biol *7*, 458.

Mogridge, J. (2004). Using light scattering to determine the stoichiometry of protein complexes. Methods Mol. Biol. *261*, 113–118.

Morgan, D.O. (2007). The Cell Cycle: Principles of Control (London: New Science Press).

Murayama, Y., and Uhlmann, F. (2014). Biochemical reconstitution of topological DNA binding by the cohesin ring. Nature *505*, 367–371.

Musacchio, A., and Salmon, E.D. (2007). The spindle-assembly checkpoint in space and time. Nat. Rev. Mol. Cell Biol. *8*, 379–393.

Nasmyth, K. (2011). Cohesin: a catenase with separate entry and exit gates? Nat Cell Biol *13*, 1170–1177.

Nasmyth, K., and Haering, C.H. (2005). The structure and function of SMC and kleisin complexes. Annu. Rev. Biochem. *74*, 595–648.

Nishiyama, T., Ladurner, R., Schmitz, J., Kreidl, E., Schleiffer, A., Bhaskara, V., Bando, M., Shirahige, K., Hyman, A.A., Mechtler, K., et al. (2010). Sororin Mediates Sister Chromatid Cohesion by Antagonizing Wapl. Cell *143*, 737–749.

O'Shaughnessy, L., and Doyle, S. (2011). Purification of proteins from baculovirus-infected insect cells. Methods Mol. Biol. *681*, 295–309.

Oliveira, R.A., Hamilton, R.S., Pauli, A., Davis, I., and Nasmyth, K. (2010). Cohesin cleavage and Cdk inhibition trigger formation of daughter nuclei. Nat. Cell Biol. *12*, 185–192.

Panizza, S., Tanaka, T., Hochwagen, A., Eisenhaber, F., and Nasmyth, K. (2000). Pds5 cooperates with cohesin in maintaining sister chromatid cohesion. Curr. Biol. *10*, 1557–1564.

Plotkin, J.B., and Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. Nat. Rev. Genet. *12*, 32–42.

Rankin, S. (2005). Sororin, the cell cycle and sister chromatid cohesion. Cell Cycle *4*, 1039–1042.

Rankin, S., Ayad, N.G., and Kirschner, M.W. (2005). Sororin, a Substrate of the Anaphase-Promoting Complex, Is Required for Sister Chromatid Cohesion in Vertebrates. Mol. Cell *18*, 185–200.

Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., and Deber, C.M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. Proc. Natl. Acad. Sci. U. S. A. *106*, 1760–1765.

Roig, M.B., Löwe, J., Chan, K.-L., Becköuet, F., Metson, J., and Nasmyth, K. (2014). Structure and function of cohesin's Scc3/SA regulatory subunit. FEBS Lett.

Rowland, B.D., Roig, M.B., Nishino, T., Kurze, A., Uluocak, P., Mishra, A., Beckouet, F., Underwood, P., Metson, J., Imre, R., et al. (2009). Building sister chromatid cohesion: Smc3 acetylation counteracts an antiestablishment activity. Mol Cell *33*, 763–774.

Sagan, C., and Druyan, A. (1995). The Demon-Haunted World: Science As a Candle in the Dark (Paw Prints).

Sahdev, S., Khattar, S.K., and Saini, K.S. (2008). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol. Cell. Biochem. *307*, 249–264.

Saibil, H. (2013). Chaperone machines for protein folding, unfolding and disaggregation. Nat. Rev. Mol. Cell Biol. *14*, 630–642.

San-Miguel, T., Pérez-Bermúdez, P., and Gavidia, I. (2013). Production of soluble eukaryotic recombinant proteins in E. coli is favoured in early log-phase cultures induced at low temperature. Springerplus *2*, 89.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schmitz, J., Watrin, E., Lénárt, P., Mechtler, K., and Peters, J.-M. (2007). Sororin Is Required for Stable Binding of Cohesin to Chromatin and for Sister Chromatid Cohesion in Interphase. Curr. Biol. *17*, 630–636.

Schur, F.K.M., Hagen, W.J.H., de Marco, A., and Briggs, J.A.G. (2013). Determination of protein structure at 8.5Å resolution using cryo-electron tomography and sub-tomogram averaging. J. Struct. Biol. *184*, 394–400.

Shintomi, K., and Hirano, T. (2009). Releasing cohesin from chromosome arms in early mitosis: opposing actions of Wapl-Pds5 and Sgo1. Genes Dev. *23*, 2224–2236.

Singh-Blom, A., Hughes, R.A., and Ellington, A.D. (2013). Residue-Specific Incorporation of Unnatural Amino Acids into Proteins In Vitro and In Vivo. In Enzyme Engineering: Methods and Protocols, (Totowa, NJ: Humana Press), pp. 93–114.

Skube, S.B., Chaverri, J.M., and Goodson, H. V (2010). Effect of GFP tags on the localization of EB1 and EB1 fragments in vivo. Cytoskeleton (Hoboken). *67*, 1–12.

Smits, A.H., Jansen, P.W.T.C., Poser, I., Hyman, A.A., and Vermeulen, M. (2013). Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Res. *41*, e28.

Song, J. (2013). Why do proteins aggregate? “Intrinsically insoluble proteins” and “dark mediators” revealed by studies on “insoluble proteins” solubilized in pure water. F1000Research *2*, 94.

Sullivan, M., and Morgan, D.O. (2007). Finishing mitosis, one step at a time. Nat. Rev. Mol. Cell Biol. *8*, 894–903.

Sumara, I., Vorlaufer, E., Gieffers, C., Peters, B.H., and Peters, J.M. (2000). Characterization of vertebrate cohesin complexes and their regulation in prophase. J. Cell Biol. *151*, 749–762.

Sumner, A.T. (1991). Scanning electron microscopy of mammalian chromosomes from prophase to telophase. Chromosoma *100*, 410–418.

Tanaka, T., Cosma, M.P., Wirth, K., and Nasmyth, K. (1999). Identification of Cohesin Association Sites at Centromeres and along Chromosome Arms. Cell *98*, 847–858.

Tedeschi, A., Wutz, G., Huet, S., Jaritz, M., Wuensche, A., Schirghuber, E., Davidson, I.F., Tang, W., Cisneros, D.A., Bhaskara, V., et al. (2013). Wapl is an essential regulator of chromatin structure and chromosome segregation. Nature *501*, 564–568.

Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biotechnol. *60*, 523–533.

Toth, A., Ciosk, R., Uhlmann, F., Galova, M., Schleiffer, A., and Nasmyth, K. (1999). Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication. Genes Dev *13*, 320–333.

Uhlmann, F., Lottspeich, F., and Nasmyth, K. (1999). Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. Nature *400*, 37–42.

Uhlmann, F., Wernic, D., Poupart, M.-A., Koonin, E. V, and Nasmyth, K. (2000). Cleavage of Cohesin by the CD Clan Protease Separin Triggers Anaphase in Yeast. Cell *103*, 375–386.

Ünal, E., Heidinger-Pauli, J.M., Kim, W., Guacci, V., Onn, I., Gygi, S.P., and Koshland, D.E. (2008). A Molecular Determinant for the Establishment of Sister Chromatid Cohesion. Science *321*, 566–569.

Varshavsky, A. (1997). The N-end rule pathway of protein degradation. Genes Cells *2*, 13–28.

Velappan, N., Sblattero, D., Chasteen, L., Pavlik, P., and Bradbury, A.R.M. (2007). Plasmid incompatibility: more compatible than previously thought? Protein Eng. Des. Sel. *20*, 309–313.

Vera, A., González-Montalbán, N., Arís, A., and Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. Biotechnol. Bioeng. *96*, 1101–1106.

Waizenegger, I.C., Hauf, S., Meinke, A., and Peters, J.-M. (2000). Two Distinct Pathways Remove Mammalian Cohesin from Chromosome Arms in Prophase and from Centromeres in Anaphase. Cell *103*, 399–410.

Wang, S.-W., Read, R.L., and Norbury, C.J. (2002). Fission yeast Pds5 is required for accurate chromosome segregation and for survival after DNA damage or metaphase arrest. J. Cell Sci. *115*, 587–598.

Waugh, D.S. (2005). Making the most of affinity tags. Trends Biotechnol. *23*, 316–320.

Weitzer, S., Lehane, C., and Uhlmann, F. (2003). A Model for ATP Hydrolysis-Dependent Binding of Cohesin to DNA. Curr. Biol. *13*, 1930–1940.

Whelan, G., Kreidl, E., Wutz, G., Egner, A., Peters, J.-M., and Eichele, G. (2012). Cohesin acetyltransferase Esco2 is a cell viability factor and is required for cohesion in pericentric heterochromatin. EMBO J. *31*, 71–82.

Wu, F.M., Nguyen, J. V, and Rankin, S. (2011). A Conserved Motif at the C Terminus of Sororin Is Required for Sister Chromatid Cohesion. J. Biol. Chem. *286*, 3579–3586.

Yeh, E., Haase, J., Paliulis, L. V, Joglekar, A., Bond, L., Bouck, D., Salmon, E.D., and Bloom, K.S. (2008). Pericentric Chromatin Is Organized into an Intramolecular Loop in Mitosis. Curr. Biol. *18*, 81–90.

Zhang, N., and Pati, D. (2012). Sororin is a master regulator of sister chromatid cohesion and separation. Cell Cycle *11*, 2073–2083.

Appendix 1

Oligonucleotide list: cloning

Name	Sequence	Details
KK77	GGAGATATACCATGGCTACTGCTGTGCGTC GCTCAACTAG	NcoI + *SCC3* NT
KK27	ACTTGAAAGCTTTTATTTCTCAAACTGCGG ATGTGACCAATCTTGTGTGATTTCGTCGCTG TTGTCTA	HindIII + StrepII + *SCC3* CT
KK79	CAAGATTAAAAGCTTGTTTAACTTTAAGAA GGAGATATAAATATGGCTGTGACCGAAAAC	HindIII + RBS + *SCC1* NT
KK29	ACTTGACTCGAGTTAGTGGTGGTGGTGGTG ATTGCGTTGATAAAGCGTTCAAACAGTGCC GGTTTC	XhoI + 6×His + *SCC1* CT
KK93	GGAGATATACCATGGCTTACATCAAACGTG TTATCATCAAGGGCTTCAAGACCTA	NcoI + *SMC3* NT
KK23	ACTTGAAAGCTTTTATTTCTCAAACTGCGG ATGTGACCAAACTTCTGCAAATTTATTTGA ACCGC	HindIII + StrepII + *SMC3* CT
KK95	GAGAAATAAAAGCTTGTTTAACTTTAAGAA GGAGATATAAATATGGCTGGTCGCCTGGTT GGTCTGGAACTGAGCAATTT	HindIII + RBS + *SMC1* NT
KK70	GTGGTGGTGCTCGAGTTAATGATGGTGATG ATGATGTTCTGCGTAGTTGGACAGATCCAG GGT	XhoI + 6×His + *SMC1* CT
KK97	GGAGATATACCATGGCTGTGACCGAAAACC CGCAGCGTCTGACCGTTCTGCGTCT	NcoI + *SCC1* NT
KK132	TGCGGCCGCAAGCTTTTAGTGGTGGTGGTG GTGATGTGCGTTGATAAAGCGTTCAAACAG TGCCGGTTTC	HindIII + 6×His + *SCC1* CT
KK135	GTGGTGGTGCTCGAGTTATTTGTCATCGTC ATCATAATCTTCTGCGTAGTTGGACAGATCC AGGGTGATAATTTTAGAGC	XhoI + FLAG + *SMC1* CT

Oligonucleotide list: sequencing

Name	Sequence	Details
KK55	GCGGCGATGGCGGAGCTGAA	pet28a vector sequencing
KK56	TGAAGGGCAATCAGCTGTTG	
KK57	CCAGTTGTTTACCCTCACAA	
KK58	ACCCGACAGGACTATAAAGA	
KK59	GTAGCGTTGCCAATGATGTT	
KK60	AATCCCTTATAAATCAAAAG	
KK53	GCTATATCTTTTCCAGGACT	*SCC3* FWD-1
KK54	ATGTTGACGATGAGGAATTG	*SCC3* FWD-2
KK43	CCTTCACGATAATGCTTCCAAAGTG	*SCC3* FWD-3
KK61	AGTCCTGGAAAAGATATAGC	*SCC3* REV-1
KK121	GCCGGCCGGTAACATTGACA	*SCC1* FWD-1
KK46	TGTCAATGTTACCGGCCGGC	*SCC1* REV-1
KK122	CCCGATTAGTTCCGAAACCA	*SMC1* FWD-1
KK47	TGGTTTCGGAACTAATCGGG	*SMC1* REV-1
KK48	AGTTTACCGCGAATGCCCTT	*SMC1* REV-2
KK49	GCTGATTCCTTAACGAACGA	*SMC1* REV-3
KK50	AAGCCAAACTGCGTGAAAGC	*SMC3* FWD-1
KK51	AAAAGCGGTCAAGCATGTGT	*SMC3* FWD-2
KK52	AGACCATGATTAAAAAGACC	*SMC3* FWD-3
KK123	GCTTTCACGCAGTTTGGCTT	*SMC3* REV-1

Oligonucleotide list: site-directed mutagenesis

Name	Sequence	Details
KK107	AAGATtgcAAGTGGAAAAAGGGCATTCGC	Smc1 L639C
KK108	CACTTgcaATCTTTTGCGATGTTCAGGGT	
KK101	ATACGtgcGAAACCGCGACGCTGATCATG	Smc3 E570C
KK102	GTTTCgcaCGTATCCACGACAATATGGAA	
KK81	GTCGAttgCACGTTTCAGAAAGTCTCTGA	Smc3 S1043C
KK82	ACGTGcaaTCGACGGCGTTAACTTTCTGT	
KK41	AAGAAaaaCTTGGCCCGCTGTTCAAAAAA	Removal of HindIII site from *SCC3*
KK42	CCAAGtttTTCTTTGTTTAGTGCCACACG	

Appendix 2

Construct and *E. coli* expression strain list

Plasmid ID	Expression strain ID	Construct in pET28a
KKD26	KKEX14	Scc3-StrepII + Scc1-6×His
KKD24	KKEX17	Smc3-StrepII + Smc1-6×His
KKD34	KKEX19	Smc3(S1043C)-StrepII + Smc1-6×His
KKD42	KKEX20	Smc3(E570C)-StrepII + Smc1(L639C)- 6×His
KKD58	KKEX38	Scc1-6×His
KKD63	KKEX42	Scc1-6×His + Smc1-FLAG
KKD20	KKEX37	Smc3-StrepII

The pET28a vector carries the kanamycin resistance selection marker. The BL21 DE3 *E. coli* strain used for expression additionally hosts the chloramphenicol resistance gene, so amplification of bacterial cultures for protein expression was performed in media supplemented in both kanamycin and chloramphenicol.